

SYNTHESE LIBRARY / VOLUME 55

ADOLF GRÜNBAUM

PHILOSOPHICAL PROBLEMS OF SPACE AND TIME

Second, enlarged edition

BOSTON STUDIES IN THE PHILOSOPHY OF SCIENCE
VOLUME XII

Edited by Robert S. Cohen and Marx W. Wartofsky



D. REIDEL PUBLISHING COMPANY

DORDRECHT-HOLLAND / BOSTON-U.S.A.

BOSTON STUDIES IN THE
PHILOSOPHY OF SCIENCE

VOLUME XII

PHILOSOPHICAL PROBLEMS
OF SPACE AND TIME

Second, enlarged edition

By ADOLF GRÜNBAUM

University of Pittsburgh, Pittsburgh, Pa., U.S.A.

INSTITUTE OF
ADVANCED STUDY
LIBRARY SIMLA

'In the past ten years Adolf Grünbaum has written three books and many articles..., and other philosophers and scientists have argued for and against his powerful and sophisticated theses...' (TLS, June, 1973).

'The depth of Grünbaum's work is indicated by the range of philosophical topics that it covers. ...The background for the work is empiricist, but it is remarkably free of doctrinaire lines'. (Arthur Fine, *Synthese*, 22/4)

This second, enlarged edition of Grünbaum's influential work *Philosophical Problems of Space and Time* is a substantial elaboration, correction and extension of the ideas presented in the original 1963 edition, which was the first rather comprehensive treatise on its subject since Hans Reichenbach's classic *Philosophie der Raum-Zeit-Lehre* (1928). The second edition augments the original text by more than 100% new material consisting of seven Studies written during the decade 1963-1973, a new Preface, a large bibliography following each chapter, an Appendix containing detailed chapter-by-chapter commentaries, and a General Index of Subjects and Names.

The enlarged edition includes treatments of important topics not only from contemporary relativity theory but also from outside that theory. It is divided into 22 chapters with an Appendix, and these chapters fall into four major parts.

Part I, which is concerned with metrical aspects of space, time and space-time, centers its philosophical account of geometry and chronometry on the ontological status of metrical equality. The treatment is both constructive and critical. The assessment of the views of major modern writers includes, for example, a critique of Pierre Duhem's philosophy of science which is occasioned by the examination of Einstein's philosophy of geometry. And there is a chapter on how modern science avoids Zeno's metrical paradox of extension.

(continued on backflap)

D. REIDEL PUBLISHING COMPANY
DORDRECHT-HOLLAND / BOSTON-U.S.A.



Handwritten text, possibly a date or signature, is visible but illegible.

BOSTON STUDIES IN THE PHILOSOPHY OF SCIENCE

VOLUME VII

PHILOSOPHICAL PROBLEMS OF SPACE AND TIME

SYNTHESE LIBRARY

MONOGRAPHS ON EPISTEMOLOGY,
LOGIC, METHODOLOGY, PHILOSOPHY OF SCIENCE,
SOCIOLOGY OF SCIENCE AND OF KNOWLEDGE,
AND ON THE MATHEMATICAL METHODS OF
SOCIAL AND BEHAVIORAL SCIENCES

Editors:

DONALD DAVIDSON, *The Rockefeller University and Princeton University*

JAAKKO HINTIKKA, *Academy of Finland and Stanford University*

GABRIËL NUCHELMANS, *University of Leyden*

WESLEY C. SALMON, *University of Arizona*

VOLUME 55

BOSTON STUDIES IN THE PHILOSOPHY OF SCIENCE

VOLUME XII

EDITED BY ROBERT S. COHEN AND MARK W. WARTOFSKY

ADOLF GRÜNBAUM

Andrew Mellon Professor of Philosophy, University of Pittsburgh

PHILOSOPHICAL PROBLEMS OF SPACE AND TIME

Second, enlarged edition



D. REIDEL PUBLISHING COMPANY

DORDRECHT-HOLLAND / BOSTON-U.S.A.

First printing: December 1973

Second printing: 1974

Library of Congress Catalog Card Number 73 75763



U.S. Shir-la

Cloth edition: ISBN 90 277 0357 4

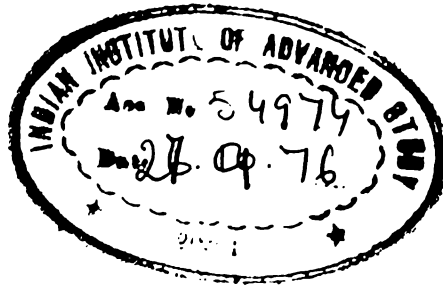
Pack edition: ISBN 90 277 0358 2



00054974

by D. Reidel Publishing Company,
P.O. Box 17, Dordrecht, Holland

Sold and distributed in the U.S.A., Canada, and Mexico
by D. Reidel Publishing Company, Inc
306 Dartmouth Street, Boston,
Mass. 02116, U.S.A.



All Rights Reserved

Copyright © 1973 by D. Reidel Publishing Company, Dordrecht, Holland

No part of this book may be reproduced in any form, by print, photoprint, microfilm,
or any other means, without written permission from the publisher

Printed in The Netherlands by D. Reidel, Dordrecht

To Thelma

EDITORIAL INTRODUCTION	XIII
PREFACE TO THE FIRST EDITION	XV
PREFACE TO THE SECOND, ENLARGED EDITION	XVII
ACKNOWLEDGMENTS	XXI

PART I PHILOSOPHICAL PROBLEMS OF THE
METRIC OF SPACE AND TIME

<i>Chapter 1.</i>	Spatial and Temporal Congruence in Physics: A Critical Comparison of the Conceptions of Newton, Riemann, Poincaré, Eddington, Bridgman, Russell, and Whitehead	3
	A. Newton	4
	B. Riemann	8
	C. Poincaré	18
	D. Eddington	24
	E. Bridgman	41
	F. Russell	44
	G. Whitehead	48
<i>Chapter 2.</i>	The Significance of Alternative Time Metrizations in Newtonian Mechanics and in the General Theory of Relativity	66
	A. Newtonian Mechanics	66
	B. The General Theory of Relativity	77
<i>Chapter 3.</i>	Critique of Reichenbach's and Carnap's Philosophy of Geometry	81
	A. The Status of "Universal Forces"	81
	B. The "Relativity of Geometry"	98
<i>Chapter 4.</i>	Critique of Einstein's Philosophy of Geometry	106
	A. An Appraisal of Duhem's Account of the Falsifi-	

	ability of Isolated Empirical Hypotheses in Its Bearing on Einstein's Conception of the Interdependence of Geometry and Physics	106
	I. The Trivial Validity of the D-Thesis	111
	II. The Untenability of the Non-Trivial D-Thesis	114
	B. The Interdependence of Geometry and Physics in Poincaré's Conventionalism	115
	C. Critical Evaluation of Einstein's Conception of the Interdependence of Geometry and Physics: Physical Geometry as a Counter-Example to the Non-Trivial D-Thesis	131
<i>Chapter 5.</i>	Empiricism and the Geometry of Visual Space	152
<i>Chapter 6.</i>	The Resolution of Zeno's Metrical Paradox of Extension for the Mathematical Continua of Space and Time	158

PART II. PHILOSOPHICAL PROBLEMS OF THE TOPOLOGY OF TIME AND SPACE

<i>Chapter 7.</i>	The Causal Theory of Time	179
	A. Closed Time	197
	B. Open Time	203
<i>Chapter 8.</i>	The Anisotropy of Time	209
	A. Is There a Thermodynamic Basis for the Anisotropy of Time?	209
	I. The Entropy Law of Classical Thermodynamics	219
	II. The Statistical Analogue of the Entropy Law	236
	B. Are There Non-Thermodynamic Foundations for the Anisotropy of Time?	264
<i>Chapter 9.</i>	The Asymmetry of Retrodictability and Predictability, the Compossibility of Explanation of the Past and Prediction of the Future, and Mechanism vs. Teleology	281
	A. The Conditions of Retrodictability and Non-Predictability	281
	B. The Physical Basis for the Anisotropy of Psychological Time	289

C	The Bearing of Retrodictability and Non-Predictability on the Compossibility of Explainability and Predictability	290
I	Evolutionary Theory	300
II	The Paresis Case	303
III	The Barometer Case	309
D	The Controversy Between Mechanism and Teleology	311
<i>Chapter 10</i>	Is There a "Flow" of Time or Temporal "Becoming"?	314
<i>Chapter 11</i>	Empiricism and the Three-Dimensionality of Space	330

PART III PHILOSOPHICAL ISSUES IN THE THEORY OF RELATIVITY

<i>Chapter 12.</i>	Philosophical Foundations of the Special Theory of Relativity, and Their Bearing on Its History	341
A.	Introduction	341
B.	Einstein's Conception of Simultaneity, Its Prevalent Misrepresentations, and Its History	342
C.	History of Einstein's Enunciation of the Limiting Character of the Velocity of Light <i>in vacuo</i>	369
D.	The Principle of the Constancy of the Speed of Light, and the Falsity of the Aether-Theoretic Lorentz-Fitzgerald Contraction Hypothesis	386
E.	The Experimental Confirmation of the Kinematics of the STR	397
F.	The Philosophical Issue Between Einstein and His Aether-Theoretic Precursors, and Its Bearing on E. T. Whittaker's History of the STR	400
<i>Chapter 13.</i>	Philosophical Appraisal of E. A. Milne's Alternative to Einstein's STR	410
<i>Chapter 14.</i>	Has the General Theory of Relativity Repudiated Absolute Space?	418
<i>Chapter 15.</i>	Philosophical Critique of Whitehead's Theory of Relativity	425
	BIBLIOGRAPHY FOR THE FIRST EDITION	429

	ability of Isolated Empirical Hypotheses in Its Bearing on Einstein's Conception of the Interdependence of Geometry and Physics	106
	I. The Trivial Validity of the D-Thesis	111
	II. The Untenability of the Non-Trivial D-Thesis	114
	B. The Interdependence of Geometry and Physics in Poincaré's Conventionalism	115
	C. Critical Evaluation of Einstein's Conception of the Interdependence of Geometry and Physics: Physical Geometry as a Counter-Example to the Non-Trivial D-Thesis	131
<i>Chapter 5.</i>	Empiricism and the Geometry of Visual Space	152
<i>Chapter 6.</i>	The Resolution of Zeno's Metrical Paradox of Extension for the Mathematical Continua of Space and Time	158

PART II. PHILOSOPHICAL PROBLEMS OF THE TOPOLOGY OF TIME AND SPACE

<i>Chapter 7.</i>	The Causal Theory of Time	179
	A. Closed Time	197
	B. Open Time	203
<i>Chapter 8.</i>	The Anisotropy of Time	209
	A. Is There a Thermodynamic Basis for the Anisotropy of Time?	209
	I. The Entropy Law of Classical Thermodynamics	219
	II. The Statistical Analogue of the Entropy Law	236
	B. Are There Non-Thermodynamic Foundations for the Anisotropy of Time?	264
<i>Chapter 9.</i>	The Asymmetry of Retrodictability and Predictability, the Compossibility of Explanation of the Past and Prediction of the Future, and Mechanism vs. Teleology	281
	A. The Conditions of Retrodictability and Non-Predictability	281
	B. The Physical Basis for the Anisotropy of Psychological Time	289

C	The Bearing of Retrodictability and Non-Predictability on the Compossibility of Explainability and Predictability	290
I	Evolutionary Theory	300
II	The Paresis Case	303
III	The Barometer Case	309
D	The Controversy Between Mechanism and Teleology	311
<i>Chapter 10</i>	Is There a "Flow" of Time or Temporal "Becoming"?	314
<i>Chapter 11</i>	Empiricism and the Three-Dimensionality of Space	330

PART III PHILOSOPHICAL ISSUES IN THE THEORY OF RELATIVITY

<i>Chapter 12</i>	Philosophical Foundations of the Special Theory of Relativity, and Their Bearing on Its History	341
A.	Introduction	341
B.	Einstein's Conception of Simultaneity, Its Prevalent Misrepresentations, and Its History	342
C.	History of Einstein's Enunciation of the Limiting Character of the Velocity of Light <i>in vacuo</i>	369
D.	The Principle of the Constancy of the Speed of Light, and the Falsity of the Aether-Theoretic Lorentz-Fitzgerald Contraction Hypothesis	386
E.	The Experimental Confirmation of the Kinematics of the STR	397
F.	The Philosophical Issue Between Einstein and His Aether-Theoretic Precursors, and Its Bearing on E. T. Whittaker's History of the STR	400
<i>Chapter 13</i>	Philosophical Appraisal of E. A. Milne's Alternative to Einstein's STR	410
<i>Chapter 14</i>	Has the General Theory of Relativity Repudiated Absolute Space?	418
<i>Chapter 15</i>	Philosophical Critique of Whitehead's Theory of Relativity	425
	BIBLIOGRAPHY FOR THE FIRST EDITION	429

PART IV. SUPPLEMENTARY STUDIES 1964-1973

1. Supplement to Part I

<i>Chapter 16.</i> Space, Time and Falsifiability (First Installment)	449
Abstract	449
Introduction	450
Criteria for Intrinsicness vs. Extrinsicness of Metrics and of Relations on Manifolds: Contents	457
1. Singly and Multiply Extended Manifolds	458
2. Intrinsicness vs. Extrinsicness of Metrics, Metrical Equalities, and Congruences	468
3. What are the Logical Connections, if any, between Alternative Metrizable, Intrinsic Metric Amor- phousness, and the Convention-ladenness of Metr- ical Comparisons?	547
4. Intrinsicness and Extrinsicness of a Relation on a Manifold	563
<i>Chapter 17.</i> Can We Ascertain the Falsity of a Scientific Hypoth- esis?	569
1. Introduction	569
2. Purported Disproofs of Hypotheses in Biology and Astronomy	572
3. Is it NEVER Possible to Falsify a Hypothesis Irrevocably?	585
<i>Chapter 18.</i> Can an Infinitude of Operations Be Performed in a Finite Time?	630

2. Supplement to Part II

<i>Chapter 19.</i> Is the Coarse-Grained Entropy of Classical Statistical Mechanics an Anthropomorphism?	646
1. Introduction	646
2. Entropy Change and Arbitrariness of the Partition- ing of Phase Space	648
3. What is the Physical Significance of the Triple Role of the Entropy for the Entropy Statistics in the Class U ?	659

4 Do the Roles of Human Decision and Ignorance Impugn the Physical Significance of the Entropy Statistics for the Class U ?	663
---	-----

3 Supplement to Part III

<i>Chapter 20</i> Simultaneity by Slow Clock Transport in the Special Theory of Relativity	666
Introduction (co-authored with Wesley C. Salmon)	666
1. Summary	670
2. Examination of Ellis and Bowman's Account of Nonstandard Signal Synchronizations	671
3. The Philosophical Status of Simultaneity by Slow Clock Transport in the Special Theory of Relativity	683
<i>Chapter 21</i> The Bearing of Philosophy on the History of the Special Theory of Relativity	709
1. History and Pedagogy of the Light Principle	711
2. Contraction and Time-Dilation Hypotheses	715
3. Summary	726
<i>Chapter 22</i> General Relativity, Geometrodynamics and Ontology	728
1. Introduction	728
2. The Philosophical Status of the Metric of Space- Time in the General Theory of Relativity	730
3. The Ontology of Empty Curved Metric Space in the Geometrodynamics of Clifford and Wheeler	750
4. The Time-Orientability of Space-Time and the 'Arrow' of Time	788
APPENDIX	804
INDEX OF PERSONAL NAMES Compiled by Mr. Theodore C. Falk	857
INDEX OF SUBJECTS Compiled by Mr. Theodore C. Falk	865

It is ten years since Adolf Grünbaum published the first edition of this book. It was promptly recognized to be one of the few major works in the philosophy of the natural sciences of this generation. In part, this is so because Grünbaum has chosen a problem basic both to philosophy and to the natural sciences — the nature of space and time; and in part, this is so because he so admirably exemplifies that Aristotelian devotion to the intimate and mutual dependence of actual science and philosophical understanding. More than this, however, the quality of his work derives from his achievement in combining detail with scope. The problems of space and time have been among the most difficult in contemporary and classical thought, and Grünbaum has been responsible to the full depth and complexity of these difficulties. This revised and enlarged second edition is a work in progress, in the tradition of reflective analysis of modern science of such figures as Ehrenfest and Reichenbach.

In publishing this work among the *Boston Studies in the Philosophy of Science*, we hope to contribute to and encourage that broad tradition of natural philosophy which is marked by the close collaboration of philosophers and scientists. To this end, we have published the proceedings of our Colloquia, of meetings and conferences here and abroad, as well as the works of single authors. From the start, Adolf Grünbaum has been among the staunchest supporters of the Boston Colloquium for the Philosophy of Science. We are pleased to be able to include his significant treatise in our series.

ROBERT S. COHEN
MARX W. WARTOFSKY

My principal intellectual debt is to Hans Reichenbach's outstanding work *Philosophie der Raum-Zeit-Lehre* (Berlin, 1928) and to A. d'Abro's remarkable *The Evolution of Scientific Thought from Newton to Einstein* (New York, 1927).

In addition to those mentioned within the text, a number of colleagues and friends offered suggestions or criticisms which were of significant benefit to me in developing the ideas of this book. Among these, I should name especially the scientists Peter Havas, Allen Janis, Samuel Gulden, E. L. Hill, and Albert Wilansky as well as the following philosophers: Henry Mehlberg, Wilfrid Sellars, Abner Shimony, Grover Maxwell, Herbert Feigl, Hilary Putnam, Paul K. Feyerabend, Ernest Nagel, Nicholas Rescher, Sidney Morgenbesser, and Robert S. Cohen. Fruitful exchanges with some of these colleagues were made possible by the stimulating sessions of the Minnesota Center for Philosophy of Science to whose Director, Professor Herbert Feigl, I am very grateful for much encouragement.

I wish to thank Mrs. Helen Farrell of Bethlehem, Pennsylvania for typing an early draft of a portion of the manuscript, and Mrs. Elizabeth McMunn whose intelligence and conscientiousness were invaluable in the preparation of the final text for the printer. I am also indebted to Mr. Richard K. Martin for assistance in the preparation of the index and in the drawing of diagrams.

I have drawn on material which appeared in earlier versions in the following prior essays of mine: "Geometry, Chronometry and Empiricism", *Minnesota Studies in the Philosophy of Science* (ed. H. Feigl and G. Maxwell), Vol. III, Minneapolis, 1962, pp. 405-526, and "Carnap's Views on the Foundations of Geometry," in: P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Open Court Publishing Company, LaSalle, 1963, pp. 599-684. I wish to thank the editors and publishers of these essays for their kind permissions.

A decade ago the first edition of this book was published by Alfred A. Knopf in New York (1963) and was also issued by Routledge and Kegan Paul in London (1964). A good many chapters of this edition were extensively revised and/or enlarged for the Russian translation *Filosofski Problemy Prostranstva I Vremeni* which appeared in the Soviet Union (1969) under the imprint of Progress Publishers, Moscow. The work had grown out of the tradition of Hans Reichenbach's philosophy of space, time and space-time as modified by *ontological* influences emanating from Riemann and Weyl.

Well before the first edition had gone out of print in 1971, I had become dissatisfied with it in at least two major ways, neither of which I had had the opportunity to remedy as part of the revisions that were incorporated in the Soviet edition. My two main dissatisfactions were due to the following: (1) The valuable contributions by my Pittsburgh colleague G. J. Massey¹ and by my former student B. van Fraassen² to 'A Panel Discussion of Grünbaum's Philosophy of Science' (*Philosophy of Science* 36 [1969], pp. 331-399) had justly called for the provision of a more precise and more coherent account than heretofore of the Riemannian conception of an intrinsic as opposed to an extrinsic metric, a conception which I invoked foundationally in the ontology of physical geometry that I had espoused in the book, and (2) I came to think that the first edition's Chapter 14 'Has the General Theory of Relativity Repudiated Absolute Space?' was altogether insufficient to assure a *balanced* allocation of space in it between the special and general theories of relativity ('STR' and 'GTR' respectively). To remedy this imbalance, I have written a lengthy new chapter devoted to the GTR for the present edition (Chapter 22). Topics directly growing out of Einstein's theory of relativity ('TR')

¹ G. J. Massey, 'Toward a Clarification of Grünbaum's Concept of Intrinsic Metric', *Philosophy of Science* 36 (1969), pp. 331-345.

² B. van Fraassen, 'On Massey's Explication of Grünbaum's Conception of Metric', *Philosophy of Science* 36 (1969), pp. 346-353.

clearly merit detailed attention in a book on the philosophy of space and time, although the second edition of the present book, no less than the first, was written with the conviction that such a book should not be confined to topics from the TR.

After the first edition had gone out of print, its contents continued to elicit considerable critical as well as supportive commentary in the literature.³ Hence I felt that it might be useful to incorporate the reprinted text of that edition within a substantially enlarged second edition. Thus, the present volume of twenty-two chapters plus *Appendix* adds further material which more than doubles the size of the first edition reproduced within its covers with the original Knopf pagination.

Chapters 16 and 22 respectively aim at now remedying the first edition's stated two major defects. Besides, all of the additional materials by which the present book augments the Knopf edition correct and extensively elaborate some of the main ideas set forth in the first edition and in later publications such as my "Reply to Hilary Putnam's 'An Examination of Grünbaum's Philosophy of Geometry'".⁴

As mentioned in the *Acknowledgments* below, the 15 chapters of the first edition as well as Chapters 16, 17, 18, and 20 were here reproduced by a photographic process. This mode of production precluded making even limited substantive corrections or addenda directly in their text. But it did prove technically feasible to correct all but three of the prior misprints which had come to the author's attention within that material.⁵ In order to compensate for the impracticability of making even small

³ See, for example, the articles in the 1972 double issue (Nos. 1/2) of *Synthese* 24 on 'Space, Time and Geometry', edited by Patrick Suppes. For earlier detailed discussions, see Bas C. van Fraassen, *An Introduction to the Philosophy of Time and Space*, Random House, New York, 1970, *passim*, and J. J. C. Smart, *Between Science and Philosophy*, Random House, New York, 1968, *passim*.

⁴ A. Grünbaum, *Geometry and Chronometry in Philosophical Perspective*, University of Minnesota Press, Minneapolis, 1968, Chapter III, and "Reply to Hilary Putnam's 'An Examination of Grünbaum's Philosophy of Geometry'", in *Boston Studies in the Philosophy of Science* (ed. by R. S. Cohen and M. W. Wartofsky), Vol. V, Reidel, Dordrecht, 1969, pp. 1-150.

⁵ The three residual misprints occur in the Bibliography of the first edition: On p. 432, Olivier Costa de Beauregard's last name should be 'Costa de Beauregard' and not merely 'de Beauregard'; on p. 438, the 4th entry under G. Ludwig ('Questions of Irreversibility and Ergodicity') should be listed as the 2nd entry under L. Rosenfeld on p. 441; the 'P. Tannery' entry on p. 444 belongs on p. 443.

changes within these chapters, I have devoted an *Appendix* at the end of the book to selected corrections and addenda. These are presented there in the form of chapter-by-chapter commentaries. Unfortunately, due to its scope alone, I was unable to react substantively to *all* of the critical and supportive discussion which the first edition, and my published elaborations of some facets of it during the intervening decade, have evoked. Therefore, in the *Appendix* I have endeavored to mitigate the incompleteness resulting from this kind of omission by also including a brief survey of some of the pertinent literature in the Commentaries, while relating the material to other relevant publications of mine as well.

Marginal notes in the text refer to the relevant section (§) of the *Appendix*. These notes, together with Mr T. C. Falk's comprehensive indices of names and subjects, should facilitate connecting the corrections, addenda and citations of omissions to the pages of the volume where attention is devoted to the same topic or to related views of the same author. As noted on p. 449 of Chapter 16, some of the critical and supportive literature not dealt with here is due to be treated in a planned Part II of my Reply to the aforementioned 'A Panel Discussion of Grünbaum's Philosophy of Science.' Part I of that Reply appears here as Chapter 16.

I am all too aware that, besides the kind of omission just mentioned, there are also more significant omissions of entire important topics whose treatment would be of interest to someone concerned with the full range of issues in space-time philosophy. Some such topics which are cognate to one or another of the chapters of this book are also briefly mentioned in the *Appendix*, along with some references to pertinent literature.

But I still adhere to my aforementioned initial conviction that essentially *confining* the book to topics directly arising from Einstein's STR and GTR would be too high a price to pay for the achievement of reasonably adequate coverage of the enormous philosophical import of the still burgeoning though troubled current TR. One, though only one, of my reasons for this conviction is set forth at the beginning of §4 of Chapter 22 by pointing to the *lacunae in the TR* which call for *non-relativistic* discussions of time's arrow, for the employment of a non-relativistic analytical mechanics of (electromagnetically) interacting particles, and for reliance on a non-relativistic quantum theory of atomic and molecular structure. I hope that the lengthy new Chapter 22 on the GTR in the present volume does redress the first edition's *imbalance* between the

amount of space devoted to it vis-à-vis the STR. But I should emphasize that the present edition makes no pretense to adequacy of coverage of the various major themes from contemporary TR. Yet, I hope that upon considering the topics from outside of the TR which I chose to treat instead, the reader may not deem this sacrifice of a certain amount of TR coverage an undue liability.

I call attention to the various acknowledgments relating to the contents of particular chapters which are made *within* these chapters, and to those in the Preface to the first edition. Hence I shall here limit myself to expressions of indebtedness which *supplement* those found elsewhere in the book.

The first 15 chapters are reproduced from the first edition, published by Alfred A. Knopf in New York in 1963.

Chapter 16 is reprinted from *Philosophy of Science* **37** (1970), pp. 469–588 and was published there as the first installment of my response to 'A Panel Discussion of Grünbaum's Philosophy of Science' by six authors, which had appeared in vol. 36 (1969), pp. 331–399.

Chapter 17 is based on a lecture which I gave at Johns Hopkins University in May 1969 as part of the first series of the Alvin and Fanny Blaustein Thalheimer Lectures. It is reprinted from a volume containing this series: M. Mandelbaum (ed.), *Observation and Theory in Science*, The Johns Hopkins Press, Baltimore, 1971, pp. 69–129.

Chapter 18 is based on a Monday Lecture delivered at the University of Chicago in April 1968 and was first published by permission of the University of Chicago in *British J. Philosophy of Science* **20** (1969), pp. 203–218.

Chapter 19 was written for a hopefully still forthcoming *Festschrift for Henry Margenau*, edited by E. Laszlo and E. B. Sellon.

Chapter 20 first appeared in *Philosophy of Science* **36** (1969), pp. 1–43 as part of 'A Panel Discussion of Simultaneity by Slow Clock Transport in the Special and General Theories of Relativity'. The chapter begins with a 4-page Introduction to this Panel Discussion, co-authored by Wesley Salmon and myself.

Chapter 21 is based on my vice-presidential address to the History and Philosophy of Science Section (L) of the American Association for the Advancement of Science, delivered during its December 1963 Cleveland meeting. It was first published in slightly different form in *Science* **143**

(1964), pp. 1406–1412, under the title ‘The Bearing of Philosophy on the History of Science’.

Chapter 22 grew out of remarks I made as session chairman of the Geometrodynamics Symposium, held at the December 1972 Boston meeting of the American Philosophical Association, Eastern Division. It was written for this book, but some excerpts from it are to appear elsewhere at about the same time by prior arrangement with the *Journal of Philosophy* and with Patrick Suppes, Editor of *Space, Time and Geometry*. Reidel.

I wish to thank the various editors, journals, and publishers as well as Wesley Salmon for their kind permissions to use materials as specified.

Grateful acknowledgment is made for permission to quote from the following authors and/or publishers:

J. L. Synge: *Relativity: The Special Theory*, North Holland Publishing Company, 1956; quotations from pp. 14, 15, and 24.

L. O’Raifeartaigh (ed.), *General Relativity*, (Oxford University Press 1972; quotations from pp. 64, 65, 68, 69, and 72 of the essay by J. Ehlers, F. A. E. Pirani and A. Schild, by permission of the Clarendon Press, Oxford.

Beginning with page by page comments on the orally delivered embryonic form of Chapter 22, John Stachel generously gave me the immense benefit of his scholarship in relativity theory during personal discussions of a good many of the main issues treated in the final version of Chapter 22. I am cordially grateful to him. My warm thanks also go to the following other colleagues and friends for their valuable assistance, either in person or by correspondence, with matters relevant to Chapter 22: Allen Janis, who was unstinting with his time, as always, when answering scientific questions and offering suggestions for improving earlier drafts; John Porter, who made himself available as a mathematical ‘resource person’ in various ways; Morris Kline, who gave me historical information regarding Riemann’s influence on Clifford; Gerald Massey, who provided illumination on issues of logical presupposition and individuation; John Winnie, who is carrying forward some of the ideas set forth in §3 (b); and Jürgen Ehlers, whose influence is evident in the chapter and who supplied some references for the Commentary on it in the *Appendix*. Appreciation is also due to Clark Glymour for correspondence clari-

tying his views and for preprints of some of the papers by him which are discussed in the chapter, and to John Earman, who—as shown by my citations of him—was as generous with sending me preprints of his work as he was taciturn about where it was to be published.

Both the *Appendix* and some of the chapters have been affected, in ways too numerous to specify, by the intellectual impact and warm friendship which I have had the good fortune to receive from my dearly valued Pittsburgh colleagues in Philosophy, especially Alan Anderson, Nuel Belnap, Richard Gale, Laurens Laudan, Gerald Massey, Storrs McCall, Nicholas Rescher, Kenneth Schaffner, and Wilfrid Sellars. I am similarly indebted and grateful to my long-time friends Wesley Salmon and Robert Cohen, as well as to Helena Eilstein and my former doctoral students J. Alberto Coffa, Philip Quinn, and Bas van Fraassen. And I had helpful correspondence with William A. Newcomb on tachyons and retrocausation.

This second edition came into being as a volume of the *Boston Studies* through the kind cooperation of Gerard F. McCauley, Gerald T. Curtis, and Anton Reidel, and at the persistent initiative of Robert S. Cohen and Marx W. Wartofsky. It was Gerard McCauley who originally encouraged me to write the first edition for publication by Knopf.

Struck by the patent inadequacies of the index in the first edition, Theodore C. Falk, who has professional competence in both philosophy and indexing, most obligingly offered to undertake the arduous labor of preparing the subject and name indices of the present edition.

Elizabeth R. McMunn bestowed the same care on the preparation of the material of this book for the press as she has on all other manuscripts of mine during the past dozen years.

I wish to thank all of these people as well as the staff of the Reidel Publishing Company very much for their protean assistance. And I am obliged to the National Science Foundation for the support of research which issued in a portion of this book over a period of time, as well as to the University of Pittsburgh for a sabbatical leave.

Last but not least, I must convey my gratitude to my distinguished Pittsburgh colleague Dr Jack D. Myers, University Professor of Medicine, who healingly wrought my almost phoenix-like recovery from a bout with insomnia and enabled me to resume work.

Pittsburgh, Pennsylvania

July 1973

PART I

Philosophical Problems of the Metric of Space and Time

Chapter I

SPATIAL AND TEMPORAL CONGRUENCE IN PHYSICS: A CRITICAL COMPARISON OF THE CONCEPTIONS OF NEWTON, RIEMANN, POINCARÉ, EDDINGTON, BRIDGMAN, RUSSELL, AND WHITEHEAD

The metrical comparisons of separate spatial and temporal intervals required for geo-chronometry involve rigid rods and isochronous clocks. Is this involvement of a transported congruence standard to which separate intervals can be referred a matter of the mere ascertainment of an otherwise intrinsic equality or inequality obtaining among these intervals? Or is reference to the congruence standard essential logically to the very existence of these relations? More specifically, we must ask the following questions:

1. What is the warrant for the claim that a solid rod remains rigid or self-congruent under transport in a spatial region free from inhomogeneous thermal, elastic, electromagnetic and other “deforming” or “perturbational” influences? The geometrically pejorative characterization of thermal and other inhomogeneities in space as “deforming” or “perturbational” is due to the fact that they issue in a dependence of the coincidence behavior of transported solid rods on the latter’s chemical composition, and *mutatis mutandis* in a like dependence of the rates of clocks.
2. What are the grounds for asserting that a clock which is not

perturbed in the sense just specified is isochronous, i.e., yields equal durations for congruent time intervals?

This pair of questions and some of their far-reaching philosophical ramifications will occupy us in this chapter. I shall endeavor to evolve answers to them in the course of a critical discussion of the relevant rival conceptions of a number of major thinkers. It will first be in Chapter Four that we shall deal with the further issues posed by the logic of making corrections to compensate for deformations and rate-variations exhibited by rods and clocks respectively when employed geo-chronometrically under perturbing conditions.

(A) NEWTON

In the *Principia*¹ Newton states his thesis of *the intrinsicity of the metric* in "container" space and the corresponding contention for absolute time as follows:

the common people conceive those quantities [i.e., time, space, place and motion] under no other notions but from the relation they bear to sensible objects. And thence arise certain prejudices, for the removing of which it will be convenient to distinguish them into absolute and relative, true and apparent, mathematical and common.² . . . because the parts of space cannot be seen, or distinguished from one another by our senses, therefore in their stead we use sensible measures of them. For from the positions and distances of things from any body considered as immovable, we define all places; and then with respect to such places, we estimate all motions, ~~considering~~ considering bodies as transferred from some of those places into others. And so, instead of absolute places and motions, we use relative ones; and that without any inconvenience in common affairs; but in philosophical disquisitions, we ought to abstract from our senses, and consider things themselves, distinct from what are only sensible measures of them. For it may be that there is no body really at rest, to which the places and motions of others may be referred³ . . . those . . . defile the purity of mathematical and

¹ I. Newton: *Principia*, edited by F. Cajori (Berkeley: University of California Press; 1947).

² *Ibid.*, p. 6.

³ *Ibid.*, p. 8.

philosophical truths, who confound real quantities with their relations and sensible measures ⁴ . . .

I. Absolute, true, and mathematical time, of itself, and from its own nature, flows equably⁵ without relation to anything external, and by another name is called duration; relative, apparent, and common time, is some sensible and external (whether accurate or unequable) measure of duration by the means of motion, which is commonly used instead of true time; such as an hour, a day, a month, a year.

II. Absolute space, in its own nature, without relation to anything external, remains always similar and immovable. Relative space is some movable dimension or measure of the absolute spaces, which our senses determine by its position to bodies; and which is commonly taken for immovable space; such is the dimension of a subterraneous, an aerial, or celestial space, determined by its position in respect of the earth. Absolute and relative space are the same in figure and magnitude; but they do not remain always numerically the same. For if the earth, for instance, moves, a space of our air, which relatively and in respect of the earth remains always the same, will at one time be one part of the absolute space into which the air passes; at another time it will be another part of the same, and so, absolutely understood, it will be continually changed.⁶ . . . Absolute time, in astronomy, is distinguished from relative, by the equation or correction of the apparent time. For the natural days are truly unequal, though they are commonly considered as equal, and used for a measure of time; astronomers correct this inequality that they may measure the celestial motions by a more accurate time. It may be, that there is no such thing as an equable motion, whereby time may be accurately measured. All motions may be accelerated and retarded, but the flowing of absolute time is not liable to any change. The duration or perseverance of the existence of things remain the same, whether the motions are swift or slow, or none at all; and therefore this duration ought to be distinguished from what are only sensible

⁴ *Ibid.*, p. 11.

⁵ It is Newton's conception of the attributes of "equable," (i.e., congruent) time-intervals which will be subjected to critical examination and found untenable in this chapter. But in Chapter Ten below, we shall give reasons for likewise rejecting Newton's view that the concept of "flow" has relevance to the time of physics, as distinct from the time of psychology.

⁶ I. Newton: *Principia*, *op. cit.*, p. 6.

measures thereof; and from which we deduce it, by means of the astronomical equation.⁷

Newton's fundamental contentions here are that (a) the identity of points in the physical container space in which bodies are located and of the instants of receptacle time at which physical events occur is autonomous and not derivative: physical things and events do not first define, by their own identity, the points and instants which constitute their loci or the loci of other things and events, and (b) receptacle space and time each have their own *intrinsic metric*, which exists quite independently of the existence of material rods and clocks in the universe, devices whose function is at best the purely epistemic one of enabling us to ascertain the intrinsic metrical relations in the receptacle space and time contingently containing them. Thus, for example, even when clocks, unlike the rotating earth, run "equably" or uniformly, these periodic devices merely record but do not first define the temporal metric. And what Newton is therefore rejecting here is a *relational* theory of space and time which asserts that (a) bodies and events first *define* (individuate) points and instants by conferring their identity upon them, thus enabling them to serve as the loci of other bodies and events, and (b) instead of having an intrinsic metric, physical space and time are metrically amorphous pending explicit or tacit appeal to the bodies which are first to define their respective metrics.

To be sure, Newton would also reject quite emphatically any identification or isomorphism of absolute space and time, on the one hand, with the psychological space and time of conscious awareness whose respective metrics are given by unaided ocular congruence and by psychological estimates of duration, on the other. But one overlooks the essential point here, if one is led to suppose with F. S. C. Northrop⁸ that the relative, apparent and common space and time which Newton contrasts with absolute, true and mathematical space and time are the private visual space and subjective psychological time of immediate sensory experience. For Newton makes it unambiguously clear, as shown

⁷ *Ibid.*, pp. 7-8.

⁸ Cf. e.g., F. S. C. Northrop: *The Meeting of East and West* (New York: The Macmillan Company; 1946), pp. 76-77.

by the quoted passages, that his relative space and time are indeed that public space and time which is defined by the system of relations between material bodies and events, and not the egocentrically private space and time of phenomenal experience. The "sensible" measures discussed by Newton as constitutive of "relative" space and time are those furnished by the public bodies of the physicist, not by the unaided ocular congruence of one's eyes or by one's mood-dependent psychological estimates of duration. This interpretation of Newton is fully attested by the following specific assertions of his:

(1) "Absolute and relative space are the same in figure and magnitude," a declaration which is incompatible with Northrop's interpretation of *relative* space as "the immediately sensed spatial extension of, and relation between, sensed data (which is a purely private space, varying with the degree of one's astigmatism or the clearness of one's vision)."⁹

(2) As examples of merely "relative" times, Newton cites any "sensible and external (whether accurate or unequable [non-uniform]) measure of duration" such as "an hour, a day, a month, a year."¹ And he adds that the apparent time commonly used as a measure of time is based on natural days which are "truly unequal," true equality being allegedly achievable by astronomical corrections compensating for the non-uniformity of the earth's rotational motion caused by tidal friction, etc.² But Northrop erroneously takes Newton's relative time to be the "immediately sensed time" which "varies from person to person, and even for a single person passes very quickly under certain circumstances and drags under others" and asserts incorrectly that Newton identified with absolute time the public time "upon which the ordinary time of social usage is based."

(3) Newton illustrates *relative* motion by reference to the kinematic relation between a body on a moving ship, the ship, and the earth, these relations being defined in the customary manner of physics without phenomenal space or time.

⁹ *Ibid.*, p. 76.

¹ I. Newton: *Principia*, *op. cit.*, p. 6.

² The logical status of the criterion of uniformity implicitly invoked here will be discussed in some detail in Chapter Two.

Northrop is entirely right in going on to say that Einstein's conceptual innovations in the theory of relativity cannot be construed, as they have been in certain untutored quarters, as the abandonment of the distinction between physically public and privately or egocentrically sensed space and time. But Northrop's misinterpretation of the Newtonian conception of "relative" space and time prevents him from pointing out that Einstein's philosophical thesis can indeed be characterized epigrammatically as the enthronement of the very relational conception of the space-time framework which Newton sought to interdict by his use of the terms "relative," "apparent," and "common" as philosophically disparaging epithets!

See Append. §1

(B) RIEMANN

Long before the theory of relativity was propounded, a relational conception of the metric of space and time diametrically opposite to Newton's was enunciated by Riemann in the following words as part of his famous "Inaugural Lecture":

Determinate parts of a manifold, distinguished by a mark or by a boundary, are called quanta. Their comparison as to quantity comes in discrete magnitudes by counting, in continuous magnitude by measurement.³ Measuring consists in superposition of the magnitudes to be compared; for measurement there is requisite some means of carrying forward one magnitude as a measure for the other. In default of this, one can compare two magnitudes only when the one is a part of the other, and even then one can only decide upon the question of more and less, not upon the question of how many. . . . in the question concerning the ultimate basis of relations of size in space. . . . the above remark is applicable, namely that while in a discrete manifold the principle of metric relations is implicit in the notion of this manifold, it must come from somewhere else in the case of a continuous manifold. Either then the actual things forming the groundwork of a space must constitute a discrete

³ Riemann apparently does not consider denumerable dense sets, which are neither discrete nor continuous, but we shall find below that his pre-Cantorean treatment of discrete and continuous types of order as *jointly exhaustive* does not vitiate the essential core of his analysis for our purposes.

manifold, or else the basis of metric relations must be sought for outside that actuality, in colligating forces that operate upon it. . . .⁴

Although we shall see below that Riemann was mistaken in supposing that the first part of this statement will bear critical scrutiny as a characterization of continuous manifolds in general, he does render here a fundamental feature of the continua of *physical space* and *time*, which are manifolds whose elements, taken singly, all have zero magnitude. And this basic feature of the spatio-temporal continua will presently be seen to invalidate decisively the Newtonian claim of the intrinsicity of the metric in empty space and time. When now proceeding to state the upshot of Riemann's declaration for the spatio-temporal congruence issue before us, we shall not need to be concerned with the inadequacies arising from Riemann's pre-Cantorean treatment of discrete and continuous types of order as *jointly exhaustive*. And we shall defer until Chapters Fourteen and Fifteen consideration of the significance of Riemann's suggestion that "the basis of metric relations must be sought for outside [the groundwork of space] . . . in colligating forces that operate upon it" for Einstein's original quest to implement Mach's Principle in the general theory of relativity.⁵

Construing Riemann's statement as applying not only to lengths but also, *mutatis mutandis*, to areas and to volumes of higher dimensions, he gives the following sufficient condition for the intrinsic definability and non-definability of a metric without claiming it to be necessary as well: in the case of a discretely ordered set, the "distance" between two elements can be defined intrinsically in a rather natural way by the cardinality of the (least) number of intervening elements.⁶ By contrast, upon

⁴ B. Riemann: "On the Hypotheses Which Lie at the Foundations of Geometry," in: David E. Smith (ed.) *A Source Book in Mathematics* (New York: Dover Publications, Inc.; 1959), Vol. II, pp. 413 and 424-25.

⁵ A. Einstein: "Prinzipielles zur allgemeinen Relativitätstheorie," *Annalen der Physik*, Vol. LV (1918), p. 241.

⁶ The *basis* for the discrete ordering is not here at issue: it can be conventional, as in the case of the letters of the alphabet, or it may arise from special properties and relations characterizing the objects possessing the specified order.

confronting the extended continuous manifolds of physical space and time—their continuity being postulated in modern physical theories apart from programs of space and time quantization—we see that neither the cardinality of intervals nor any of their other topological properties provide a basis for an *intrinsically* defined metric. The first part of this conclusion was tellingly emphasized by Cantor's proof of the equi-cardinality of all positive intervals independently of their length. Thus, there is no *intrinsic* attribute of the space between the endpoints of a line-segment AB, or any relation between these two points themselves, in virtue of which the interval AB could be said to contain the same amount of space as the space between the termini of another interval CD not coinciding with AB. Corresponding remarks apply to the time continuum. Accordingly, the continuity we postulate for physical space and time furnishes a sufficient condition for their intrinsic metrical amorphousness. And in this sense then metric geometry is concerned not with space itself but with the relations between bodies.

Clearly, this does not preclude the existence of sufficient conditions *other than continuity* for the intrinsic metrical amorphousness of sets. But one cannot invoke densely ordered, *denumerable* sets of points (instants) in an endeavor to show that discontinuous sets of such elements may likewise lack an intrinsic metric: as we shall see in Chapter Six, even without measure theory ordinary analytic geometry allows the deduction that the length of a *denumerably* infinite point set is intrinsically zero. More generally, the measure of a denumerable point set is always zero,⁷ unless one succeeds in developing a very restrictive intuitionistic measure theory of some sort.

These considerations show incidentally that space-intervals cannot be held to be merely *denumerable* aggregates within the context of the usual mathematical theory. Hence in the context of our post-Cantorean meaning of "continuous," it is actually not as damaging to Riemann's statement as it might seem *prima facie* that he neglected the *denumerable* dense sets by incorrectly treating the discrete and continuous types of order as *jointly exhaustive*. Moreover, since the distinction between denumer-

This paragraph is supplanted by pp. 534–535 of ch. 16 and by §10 in the Appendix

⁷ E. W. Hobson: *The Theory of Functions of a Real Variable* (New York: Dover Publications, Inc.; 1957), Vol. I, p. 166.

able and super-denumerable dense sets was almost certainly unknown to Riemann, it is likely that by "continuous" he merely intended the property which we now call "dense." Evidence of such an earlier usage of "continuous" is found as late as 1914.⁸

The intrinsic metric amorphousness of the spatial continuum is made further evident by reference to the axioms for spatial congruence *after* it has been specified that they are to be given a *spatial* interpretation in terms of intervals of physical space.⁹ These axioms pre-empt "congruent" (for intervals) to be a *spatial equality predicate* by assuring the reflexivity, symmetry, and transitivity of the congruence relation in the class of spatial intervals. But although having thus pre-empted the use of "congruent" and no longer being an uninterpreted axiom system, the congruence axioms still allow an *infinitude* of *mutually exclusive* congruence classes of spatial intervals, where it is to be understood that any *particular* congruence class is a *class of classes* of congruent intervals whose lengths are specified by a *particular* distance function $ds^2 = g_{ik} dx^i dx^k$. And we just saw that there are no intrinsic metric attributes of intervals which could be invoked to single out one of these congruence classes as unique.

How then can we speak of the assumedly continuous physical space as having a metric or *mutatis mutandis* suppose that the physical time continuum has a unique metric? The answer can be none other than the following:¹ Only the choice of a particular congruence standard which is *extrinsic* to the continuum itself can determine a unique congruence class, the *rigidity* or self-congruence of that standard under transport being *decreed by convention*, and similarly for the periodic devices which are

⁸ B. Russell: *Our Knowledge of the External World* (London: George Allen and Unwin, Ltd.; 1926), p. 138.

⁹ For a statement of these axioms, see A. N. Whitehead: *The Principle of Relativity* (Cambridge: Cambridge University Press; 1922), Chap. iii, pp. 42-50.

¹ The conclusion which is about to be stated will appear unfounded to those who follow A. N. Whitehead in rejecting the "bifurcation of nature," which is assumed in its premises. But later in this chapter, the reader will find a detailed rebuttal of the Whiteheadian contention that *perceptual* space and time do have an intrinsic metric such that once the allegedly illegitimate distinction between physical and perceptual space (or time) has been jettisoned, an intrinsic metric can hence be *meaningfully imputed* to physical space and time.

held to be *isochronous* (uniform) clocks. Thus the role of the spatial or temporal congruence standard cannot be construed with Newton or Russell² to be the mere ascertainment of an otherwise intrinsic equality obtaining between the intervals belonging to the congruence class specified by it. Unless one of two segments is a subset of the other, the obtaining of the congruence relation between two segments is a matter of convention, stipulation, or definition and not a factual matter concerning which empirical findings could show one to have been mistaken. And hence there can be no question at all of an empirically or factually determinate metric geometry or chronometry until after a physical stipulation of congruence.³

In the case of geometry, the specification of the intervals which are stipulated to be congruent is given by the distance function $ds = \sqrt{g_{ik}dx^i dx^k}$, congruent intervals being those which are assigned equal lengths ds by this function. Whether the intervals defined by the coincidence behavior of a transported rod not subject to "deforming influences" are those to which the distance function assigns equal lengths ds or not will depend on our selection of the functions g_{ik} . Thus, if the components of the metric tensor g_{ik} are suitably chosen in any given coordinate system, then the transported rod will have been stipulated to

² Cf. B. Russell: "Sur les Axiomes de la Géométrie," *Revue de Métaphysique et de Morale*, Vol. VII (1899), pp. 684-707.

³ A. d'Abro (A. d'Abro: *The Evolution of Scientific Thought from Newton to Einstein* [New York: Dover Publications, Inc.; 1950], p. 27) has offered an unsound illustration of the thesis that the metric in a continuum is conventional: he considers a stream of sounds of varying pitch and points out that a congruence criterion based on the successive auditory octaves of a given musical note would be at variance with the congruence defined by equal differences between the associated frequencies of vibration, since the frequency differences between successive octaves are not equal. But instead of constituting an example of the alternative metrizable of the same mathematically continuous manifold of elements, d'Abro's illustration involves the metrizations of two different manifolds only one of which is continuous in the mathematical sense. For the auditory contents sustaining the relation of being octaves of one another are elements of a merely *sensory* "continuum." Moreover, we shall see later in this chapter that while holding for the mathematical continua of physical space and time, whose elements (points and instants) are respectively alike both qualitatively and in magnitude, the thesis of the conventionality of the metric cannot be upheld for all kinds of mathematical continua, Riemann and d'Abro to the contrary notwithstanding.

be congruent to itself everywhere independently of its position and orientation. Contrariwise, by an appropriately different choice of the functions g_{ik} , the length ds of the transported rod will be made to vary with position or orientation instead of being constant. Once congruence has been specified via the distance function ds , the geodesics (straight lines)⁴ associated with that choice of congruence are determined, since the family of geodesics is defined by the variational requirement $\delta \int ds = 0$, which takes the form of a differential equation whose solution is the equation of the family of geodesics.⁵ The geometry char-

⁴ The geodesics are called "straight lines" when discussing their relations in the context of synthetic geometry. But this identification must not be taken to entail that on surfaces other than the Euclidean plane, every geodesic connection between any two points is a line of shortest distance between them. For once we abandon the restriction to Euclidean geometry, being a geodesic connection is only a necessary and not also a sufficient condition for being the shortest distance: on the surface of a sphere, for example, "it is true that the shortest distance between two points P and Q on a sphere is along a geodesic, which on the sphere is a great circle. But there are two arcs of a great circle between two of their points, and only one of them is the curve of shortest distance, except when P and Q are at the end points of a diameter, when both arcs have the same length. This example of the sphere also shows that it is not always true that through two points only one geodesic passes: when P and Q are the endpoints of a diameter any great circle through P and Q is a geodesic and a solution of the problem of finding the shortest distance between two points." Cf. D. J. Struik: *Classical Differential Geometry* (Cambridge: Addison-Wesley Publishing Company, Inc.; 1950), p. 140. It is the case, however, that "If two points in a surface are such that only one geodesic passes through them, the length of the segment of the geodesic is the shortest distance in the surface between the two points." (L. P. Eisenhart: *An Introduction to Differential Geometry* [Princeton: Princeton University Press; 1947], Sec. 32, p. 175.)

For an account of sufficient conditions for a geodesic connection being a minimum or shortest distance, see O. Bolza: *Lectures on the Calculus of Variations* (New York: G. E. Stechert; 1946), Chap. III, §§ 17–23 inclusive, and N. I. Akhiezer: *The Calculus of Variations* (New York: Blaisdell Publishing Company, 1962), Sections 3, 4, and 15.

⁵ In the *differential* calculus, there is the problem of finding maxima and minima (extrema) of a function $y = f(x)$. And a necessary condition for the occurrence of an extremum at $x = a$ is that $dy/dx = 0$ or $dy = 0$ at $x = a$. Now, in our context the *calculus of variations* deals with the following similar but more complicated problem: Find the function $y = f(x)$ —this equation representing a family of geodesic paths here—such that the definite integral taken over the function $\int ds$ of the function $y = f(x)$ shall be a minimum or (relative) maximum, i.e., an extremum for small variations which vanish at the limits of integration. We regard $\int ds$ as a function of the function $y = f(x)$, since the former will depend on the particular kind

acterizing the relations of the geodesics in question is likewise determined by the distance function ds , because the Gaussian curvature K of every surface element at any point in space is fixed by the functions g_{ik} ingredient in the distance function ds .

There are therefore alternative metrizations of the same factual coincidence relations sustained by a transported rod, and some of these alternative definitions of congruence will give rise to different metric geometries than others. Accordingly, via an appropriate definition of congruence we are free to choose as the description of a given body of spatial facts any metric geometry compatible with the existing topology. Moreover, we shall find in Chapter Three, Section B that there are infinitely many incompatible definitions of congruence which will implement the choice of any one metric geometry, be it the Euclidean one or one of the non-Euclidean geometries.

We have been speaking of alternative "definitions" of congruence. In particular, we shall have occasion to refer to one of these "definitions" as being given by unperturbed solid rods and hence as the "customary definition" of congruence. But it has been objected that such concepts as the customary concept of spatial congruence are "multiple-criterion" concepts as opposed

of path $y = f(x)$ over which the integral is taken. In analogy to the condition $dy = 0$ for an extremum in the differential calculus, the defining condition for the family of geodesics in the calculus of variations is $\delta \int ds = 0$. Expressing ds as $I dx$, it is shown in the calculus of variations (cf. H. Margenau and G. M. Murphy: *The Mathematics of Physics and Chemistry* [New York: D. Van Nostrand Company; 1943], pp. 193-95) that this defining condition requires the following differential equation, which is known as Euler's equation, to be satisfied:

$$\frac{\partial I}{\partial y} - \frac{d}{dx} \frac{\partial I}{\partial \left(\frac{dy}{dx}\right)} = 0,$$

where the symbol " ∂ " refers to partial differentiation as distinct from the variational symbol " δ ."

As a simple illustration, consider the problem of finding the geodesics of the Euclidean plane that are associated with the standard metric

$ds^2 = dx^2 + dy^2$, which can be written as $ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \equiv I dx$.

If $\int_a^1 ds$ is to be a minimum, Euler's equation must be satisfied for the case

of $I \equiv \sqrt{1 + \left(\frac{dy}{dx}\right)^2}$. Hence we have $dy/dx = m$, where m is a constant, or $y = mx + b$. As expected, this is the equation of a family of straight lines.

to "single-criterion" concepts: congruent space intervals in inertial systems could be "defined," for example, as those for which light rays require equal (one-way or round-trip) transit times no less than they can be "defined" as the intervals given by the coincidence behavior of unperturbed transported solid rods. Thus, it has been objected that it is unsound logically to speak of giving a "definition" of congruence in the spirit of Reichenbach's "coordinative definition," since no one physical criterion, such as the one based on the solid rod, can exhaustively render the actual and potential physical meaning of the concept of spatial congruence in physics. But this objection overlooks that our reference to one or another "definition" of congruence within a set of *mutually-exclusive* "definitions" of congruence does not commit us at all to the crude operationist claim that any particular "definition" chosen by the physicist exhaustively renders *the* meaning of spatial congruence in physical theory. For our concern is with the alternative metrizableability of the spatial continuum noted by Riemann and with the resulting conventionality of congruence. Hence when we speak of "defining" congruence in this context, all that we understand by a "definition" is a specification which employs one or another criterion to single out one particular congruence class from among an infinitude of mutually-exclusive congruence classes. Thus, we shall speak of a "definition" of congruence throughout this book entirely without prejudice to the fact that spatial congruence is an open multiple-criterion concept in physics in the following sense: there is a potentially growing multiplicity of compatible physical criteria rather than only one criterion by which any one spatial congruence class (e.g., the one familiar from elementary physics) can be specified to the exclusion of every other congruence class. And it is now evident that much as attention to the multiple-criterion character of concepts in physics may be philosophically salutary in other contexts, it constitutes an intrusion of a pedantic irrelevancy in the consideration of the consequences of alternative metrizableability.

See Append §3

In pointing out earlier in this section that the status of spatial

⁶ Cf. H. Putnam: "The Analytic and the Synthetic," in: H. Feigl and G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science* (Minneapolis: University of Minnesota Press; 1962), Vol. III, esp. pp. 376-81.

and temporal congruence is decisively illuminated by Riemann's theory of continuous manifolds. I stated that this theory will not bear critical scrutiny as a characterization of continuous manifolds in general. To justify and clarify this indictment, we shall now see that continuity cannot be held with Riemann to furnish a sufficient condition for the intrinsic metric amorphousness of any manifold *independently of the character of its elements*. For, as Russell saw correctly,⁷ there are continuous manifolds, such as that of colors (in the physicist's sense of spectral frequencies), in which the individual elements differ qualitatively from one another and have inherent magnitude, thus allowing for metrical comparison of the elements themselves. By contrast, in the continuous manifolds of *space* and of *time*, neither points nor instants have any inherent magnitude allowing an individual metrical comparison between them, since all points are alike, and similarly for instants. Hence in these manifolds metrical comparisons can be effected only among the *intervals* between the elements, not among the homogeneous elements themselves. And the continuity of these manifolds then assures the non-intrinsicity of the metric for their intervals.

To exhibit further the bearing of the character of the elements of a continuous manifold on the feasibility of an intrinsic metric in it, I shall contrast the status of the metric in space and time, on the one hand, with its status in both (a) the continuum of real numbers, arranged according to magnitude and (b) the quasi-continuum of masses, mass being assumed to be a property of bodies in the Newtonian sense clarified by Mach's definition.⁸

The assignment of real numbers to points of physical space in the manner of the introduction of generalized curvilinear coordinates effects only a *coordinatization* but *not* a *metrization* of the manifold of physical space. No informative metrical comparison among individual points could be made by comparing the magnitudes of their real number coordinate-names. However, within the continuous manifold formed by the real numbers

⁷ B. Russell: *The Foundations of Geometry* (New York: Dover Publications, Inc.; 1956), Secs. 63 and 64.

⁸ For a concise account of that definition, cf. L. Page: *Introduction to Theoretical Physics* (New York: D. Van Nostrand Company, Inc.; 1935), pp. 56-58.

themselves, when arranged according to magnitude, every real number is singly distinguished from and metrically comparable to every other by its inherent magnitude. And the measurement of mass can be seen to constitute a counter-example to Riemann's metrical philosophy from the following considerations.

In the Machian definition of Newtonian (gravitational and inertial) mass, the *mass ratio* of a particle B to a standard particle A is given by the magnitude ratio of the acceleration of A due to B to the acceleration of B due to A . Once the space-time metric and thereby the accelerations are fixed in the customary way, this ratio for any particular body B is *independent*, among other things, of how far apart B and A may be during their interaction. Accordingly, any affirmations of the mass equality (mass-"congruence") or inequality of two bodies will hold independently of the extent of their spatial separation. Now, the set of medium-sized bodies form a *quasi-continuum* with respect to the dyadic relations "being more massive than," and "having the same mass," i.e., they form an array which is a continuum except for the fact that several bodies can occupy the same place in the array by sustaining the mass-"congruence" relation to each other. Without having such a relation of mass-equality *ab initio*, the set of bodies do not even form a quasi-continuum. We complete the metrization of this quasi-continuum by choosing a unit of mass (e.g., one gram) and by availing ourselves of the numerical mass-ratios obtained by experiment. There is no question here of the lack of an intrinsic metric in the sense of a choice of making the mass *difference* between a given pair of bodies equal to that of another pair or not. In the resulting continuum of real mass numbers, the elements themselves have inherent magnitude and can hence be compared individually, thus defining an intrinsic metric. Unlike the point-elements of space, the elements of the set of bodies are not all alike mass-wise, and hence the metrization of the quasi-continuum which they form with respect to the relations of being more massive and having the same mass can take the form of directly comparing the individual elements of that quasi-continuum rather than only intervals between them.

If one did wish for a spatial (or temporal) analogue of the metrization of masses, one should take as the set to be metrized

not the continuum of points (or instants) but the quasi-continuum of all spatial (or temporal) *intervals*. To have used such intervals as the elements of the set to be metrized, we must have had a prior criterion of spatial congruence and of "being longer than" in order to arrange the intervals in a quasi-continuum which can then be metrized by the assignment of length numbers. This metrization would be the space or time analogue of the metrization of masses.

(C) POINCARÉ

An illustration will serve to give concrete meaning to the general formulation of the conventionality of spatial congruence which we presented in Section B as the direct outgrowth of Riemann's analysis of the status of the metric in the spatial continuum. Consider a physical surface such as part or all of an infinite blackboard and suppose it to be equipped with a network of Cartesian coordinates. The customary metrization of such a surface is based on the congruence defined by the coincidence behavior of transported rods: line segments whose termini have coordinate differences dx and dy are assigned a length ds given by $ds = \sqrt{dx^2 + dy^2}$, and the geometry associated with this metrization of the surface is, of course, Euclidean. But we are also at liberty to employ a different metrization in part or all of this space. Thus, for example, we could equally legitimately metrize the portion *above* the x -axis on our blackboard by means of the new metric $ds = \sqrt{\frac{dx^2 + dy^2}{y^2}}$. This alternative metrization is incompatible with the customary one: for example, it makes the lengths $ds = dx/y$ of horizontal segments whose termini have the same coordinate differences dx depend on *where* they are along the y -axis. Consequently, the new metric would commit us to regard a segment for which $dx = 2$ at $y = 2$ as *congruent* to a segment for which $dx = 1$ at $y = 1$, although the customary metrization would regard the length ratio of these segments to be 2:1. But, of course, the new metric does not say that a transported solid rod will coincide successively with the intervals belonging to the congruence class defined by that metric; instead it allows for this *non*-coincidence by making the length of the rod a suitably non-constant function of its position: lying parallel

to the x-axis at $y = 2$, the rod would now be assigned only one half of the length which is ascribed to it in a like orientation at $y = 1$.

Since our new metric, which was introduced by Poincaré, yields a congruence class of *line segments* different from the customary one, one wonders whether it also issues in a non-customary congruence class of *angles*. To deal with this question we note the following requisite mathematical results. The angle θ determined by the directions A and B in a Riemann space, which are defined by the displacements A^1 and B^1 respectively, is defined by the equation

$$\cos \theta = \frac{g_{1j} A^1 B^j}{\sqrt{g_{ab} A^a A^b} \cdot \sqrt{g_{ab} B^a B^b}},$$

where the g_{1j} represent the metric coefficients⁹ of the metric $ds^2 = g_{1j} dx^1 dx^j$ for line segments. Now introduce a *new* metric which has the property that its metric coefficients g°_{1j} are related to the original coefficients g_{1j} by the following so-called "conformal" transformation: $g^{\circ}_{1j} = f(x^1) g_{1j}$, where $f(x^1)$ is an analytic multiplying function of the coordinates x^1 . It is then evident from the above expression for $\cos \theta$ that the angles θ and hence the *congruence* relations among angles will be left *unchanged* by any new metrization whose metric coefficients g°_{1j} are related by a conformal transformation to the coefficients of the original metric.¹

This result enables us to see that Poincaré's metric

$$ds^2 = \frac{1}{y^2} (dx^2 + dy^2)$$

yields the *same* congruence class of *angles* as the original standard metric $ds^2 = dx^2 + dy^2$: the coefficients of Poincaré's metric are related to those of the standard metric by the multiplying function $1/y^2$.

To determine what particular paths in the semi-blackboard are the *geodesics* of our non-standard Poincaré metric

$$ds = \frac{\sqrt{dx^2 + dy^2}}{y}$$

⁹ Cf. T. Y. Thomas: *The Differential Invariants of Generalized Spaces* (Cambridge: Cambridge University Press; 1934), p. 12.

¹ *Ibid.*, p. 20.

one must substitute this ds and carry out the variation in the defining condition $\delta(ds = 0)$ for the family of geodesics. As will be recalled from footnote 5, page 14 in Section B above, the desired geodesics of our new metric must therefore be given by Euler's equation with

$$I \equiv \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{y}.$$

Upon substituting this value of I in Euler's equation, we obtain the differential equation of the family of geodesics:

$$\frac{d^2y}{dx^2} + \frac{1}{y} \left[1 + \left(\frac{dy}{dx}\right)^2 \right] = 0.$$

The solution of this equation is of the form

$$(x - k)^2 + y^2 = R^2,$$

where k and R are constants of integration. This solution represents the family of *straights* (geodesics) associated with Poincaré's metric, but—in the Euclidean language appropriate to the standard Cartesian metric—it represents a family of "circles" centered on and perpendicular to the x -axis, the upper "semi-circles" being the *geodesics* of Poincaré's remetrized half-plane above the x -axis.

The reader can convince himself that Poincaré's metrization issues in a *hyperbolic non-Euclidean* geometry on the semi-blackboard by using the new metric coefficients $g_{11} = 1/y^2$, and $g_{22} = 1/y^2$ to obtain a *negative* value of the Gaussian curvature K via Gauss's formula.² That Poincaré's metric confers a hyperbolic geometry on the very same semi-blackboard $y > 0$ which is a *Euclidean* plane relatively to the standard metrization becomes palpable geometrically upon noting that the new geodesics of Poincaré's half plane have the following properties: First, their infinitude is assured by the behavior of Poincaré's metric as $y \rightarrow 0$, and second, they satisfy the hyperbolic parallel postulate that there exists more than one coplanar parallel

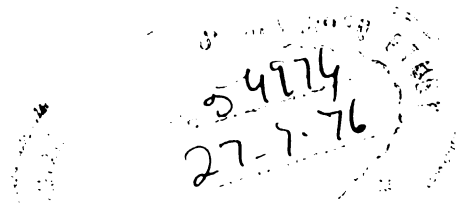
² For a statement of this formula, see, for example, F. Klein: *Vorlesungen über Nicht-Euklidische Geometrie* (Berlin: Springer-Verlag; 1928), p. 281.

straight line, since they *also* qualify as Euclidean semicircles and hence exemplify the Euclidean attribute that through a point outside a semicircle more than one semicircle can be drawn not intersecting the given semicircle.

It is apparent that the supplanting of the standard Cartesian metric by that of Poincaré has had the effect of *renaming* various paths on the semi-blackboard such that the language of hyperbolic geometry describes the very same facts of coincidence of a rod under transport on the semi-blackboard which are customarily rendered in the language of Euclidean geometry. And in the light of Riemann's searching account of spatial (and temporal) metrics, we must conclude that the hyperbolic metrization of the semi-blackboard possesses not only mathematical but also philosophical credentials as good as those of the Euclidean one.

Yet it might be objected that although non-standard metrics are legitimate philosophically, there is a pedantic artificiality and even perverse complexity in all congruence definitions which do not assign equal lengths ds to the intervals defined by the coincidence behavior of an unperturbed solid rod. The grounds of this objection would be that (a) there are no convenient and familiar natural objects whose coincidence behavior under transport furnishes a physical realization of the bizarre, non-customary congruences, and (b) after correcting for the chemically dependent distortional idiosyncrasies of various kinds of solids in inhomogeneous thermal, electric, and other fields, all transported solid bodies furnish the same physical intervals, thereby realizing one of the infinitude of incompatible mathematical congruences. *Mutatis mutandis*, the same objection might be raised to any definition of temporal congruence which is non-standard in virtue of not according with the cycles of standard unperturbed material clocks. The reply to this criticism is twofold:

(1) The *prima facie* plausibility of the demand for *simplicity* in the choice of the congruence definition gives way to second thoughts the moment it is realized that the desideratum of simplicity requires consideration not only of the congruence definition but also of the latter's bearing on the form of the associated system of geometry and physics. And our discussions in Chapters



Two and Four will show that a bizarre definition of congruence may well have to be countenanced as the price for the attainment of the over-all simplicity of the total theory. Specifically, we anticipate Chapter Two by just mentioning here that although Einstein merely alludes to the possibility of a non-customary definition of *spatial* congruence in the general theory of relativity without actually availing himself of it, he does indeed utilize in that theory what our putative objector deems a highly artificial definition of *temporal* congruence, since it is not given by the cycles of standard material clocks;

(2) It is particularly instructive to note that the cosmology of E. A. Milne⁴ postulates the actual existence in nature of two metrically different kinds of clocks whose respective periods constitute physical realizations of *incompatible* mathematical congruences. Specifically, Milne's assumptions lead to the result that there is a *non-linear* relation

$$\tau = t_0 \log \left(\frac{t}{t_0} \right) + t_0$$

between the time τ defined by periodic astronomical processes and the time t defined by atomic ones, t_0 being an appropriately chosen arbitrary constant. The non-linearity of the relation between these two kinds of time is of paramount importance here, because it assures that two intervals which are congruent in one of these two time-scales will be incongruent in the other, as is evident from the fact that the derivative $\frac{d\tau}{dt}$ is not a constant.

We can visualize the relation of the two time-scales geometrically as follows: let a half open line have the t -scale point $t = 0$ as an extremity, and use *equal* spatial intervals on the line to represent equal time intervals of the t -scale. Then the τ -scale would be represented on this same line by metrizing it so as to chop it up into spatial intervals that get progressively *shorter* in the direction of the point $t = 0$, a point which does not, how-

³ A. Einstein: "The Foundations of the General Theory of Relativity," *The Principle of Relativity, A Collection of Original Memoirs* (New York: Dover Publications, Inc.; 1952), p. 161.

⁴ E. A. Milne: *Kinematic Relativity* (Oxford: Oxford University Press; 1948), p. 22.

ever, belong to the τ -scale since $\tau \rightarrow -\infty$ as $t \rightarrow 0$. Thus, in the direction of $t = 0$ (toward the past), equal time intervals on the τ -scale correspond to ever smaller intervals on the t -scale.

Clearly, it would be utterly gratuitous to regard one of Milne's two congruences as bizarre, since each of them is presumed to have a physical realization. And the choice between these scales is incontestably conventional, for it is made quite clear in Milne's theory that their associated different metric descriptions of the world are factually equivalent and hence equally true.

What would be the verdict of the Newtonian proponent of the intrinsicity of the metric on the examples of alternative metrization which we gave both for space (Poincaré's hyperbolic metrization of the half-plane) and also for time (general theory of relativity and Milne's cosmology)? He would first note correctly that once it is understood that the term "congruent," as applied to intervals, is to denote a reflexive, symmetrical and transitive relation in this class of geometrical configurations, then the use of this term is restricted to designating a spatial equality relation. But then the Newtonian would proceed to claim unjustifiably that the spatial equality obtaining between congruent line segments of physical space (or between regions of surfaces and of 3-space respectively) consists in their each containing *the same intrinsic amount of space*. And having introduced this false premise, he would feel entitled to contend that first, it is *never* legitimate to choose arbitrarily what specific intervals are to be regarded as congruent, and second, as a corollary of this lack of choice, there is no room for selecting the lines which are to be regarded as straight and hence no choice among alternative geometric descriptions of actual physical space for the following reason: the geodesic requirement $\delta \int ds = 0$ must be satisfied by the straight lines subject to the restriction that only the members of the unique class of *intrinsically equal* line segments may ever be assigned the same length ds . By the same token, the Newtonian asserts that only "truly" (intrinsically) equal time intervals may be regarded as congruent, and he therefore holds that there exists only one admissible congruence class in the time continuum, a conclusion which he then attempts to buttress further by adducing certain causal considerations from Newtonian dynamics which will be refuted in Chapter Two below.

(D) EDDINGTON

Poincaré's view that the epistemological status of *congruence* is pivotal for the philosophical assessment of the issue of Euclideanism vs. non-Euclideanism has been rejected by Eddington. According to Eddington, the thesis that congruence (for line segments or time intervals) is conventional is true only in the trivial sense that "the meaning of every word in the language is conventional."⁵ Commenting on Poincaré's statement that we can always avail ourselves of alternative metrizable to give a Euclidean interpretation of any results of stellar parallax measurements⁶ Eddington writes:

Poincaré's brilliant exposition is a great help in understanding the problem now confronting us. He brings out the interdependence between geometrical laws and physical laws, which we have to bear in mind continually.⁷ We can add on to one set of laws that which we subtract from the other set. I admit that space is conventional—for that matter, the meaning of every word in the language is conventional. Moreover, we have actually arrived at the parting of the ways imagined by Poincaré, though the crucial experiment is not precisely the one he mentions. But I deliberately adopt the alternative, which, he takes for granted, everyone would consider less advantageous. I call the space thus chosen *physical space*, and its geometry *natural geometry*, thus admitting that other conventional meanings of space and geometry are possible. If it were only a question of the meaning of space—a rather vague term—these other possibilities might have some advantages. But the meaning assigned to length and distance has to go along with the meaning assigned to space. Now these are quantities which the physicist has been accustomed to measure with great accuracy; and they enter fundamentally into the whole of our experimental knowledge of the world. . . . Are we to be robbed of the terms in which we are accustomed to describe that knowledge?⁸

⁵ A. S. Eddington: *Space, Time and Gravitation* (Cambridge: Cambridge University Press; 1953), p. 9.

⁶ Cf. H. Poincaré: *The Foundations of Science* (Lancaster: The Science Press; 1946), p. 81, and the discussion in Chapter Four below.

⁷ This interdependence will be analyzed in Chapter Four below.

⁸ A. S. Eddington: *Space, Time and Gravitation*, *op. cit.*, pp. 9–10.

Eddington maintains that instead of being an insight into the status of spatial or temporal equality, the conventionality of congruence is a semantical platitude expressing our freedom to decree the referents of the word "congruent," a freedom which we can exercise in regard to any linguistic symbols whatever which have not already been pre-empted semantically. Thus, we are being told that though the conventionality of congruence is merely an unenlightening triviality holding for the language of any field of inquiry whatever, it has been misleadingly inflated into a philosophical doctrine about the relation of spatio-temporal equality purporting to codify fundamental features endemic to the materials of geo-chronometry. In particular, Eddington objects to Poincaré's willingness to guarantee the retention of Euclidean geometry by resorting to an alternative metrization: in the context of general relativity, the retention of Euclideanism would indeed require a congruence definition different from the customary one, as we shall see in Chapter Three. Regarding the possibility of a remetrizational retention of Euclidean geometry as merely illustrative of being able to avail oneself of the conventionality of all language, Eddington would rule out such a procedure on the grounds that the customary definition of spatial congruence which would be supplanted by the remetrization retains its usefulness.

Eddington's conclusion that only the use of the word "congruent" but not the ascription of the congruence *relation* can be held to be a matter of convention has also been defended by a cognate argument which invokes the theory of models of uninterpreted formal calculi as follows: (1) physical geometry is a spatially interpreted abstract calculus, and this interpretation of a formal system was effected by semantical rules which are all equally conventional and among which the definition of the relation term "congruent" (for line segments) does *not* occupy an epistemologically distinguished position, since we are just as free to give a *non*-customary interpretation of the abstract sign "point" as of the sign "congruent"; (2) this model-theoretic conception makes it apparent that there can be no basis at all for an epistemological distinction *within* the system of *physical* geo-chronometry between factual statements, on the one hand, and

supposedly conventional assertions of rigidity and isochronism on the other; (3) the factual credentials of physical geometry or chronometry can no more be impugned by adducing the alleged conventionality of rigidity and isochronism than one could gainsay the factuality of genetics by incorrectly affirming the conventionality of the relation of uniting which obtains between two gametes when forming a zygote.

When defending the alternative metrizable of space and time and the resulting possibility of giving either a Euclidean or a non-Euclidean description of the same spatial facts, Poincaré had construed the conventionality of congruence as an epistemological discovery about the status of the relation of spatial or temporal equality. The proponent of the foregoing model-theoretic argument therefore indicts Poincaré's defense of the feasibility of *choosing* the metric geometry as amiss, misleading, and unnecessary, deeming this choice to be automatically assured by the theory of models. And, by the same token, this critic maintains that there is no more reason for Poincaré's inquiry into the *empirical* credentials of metric geometry as such than there would be for instituting a philosophical inquiry as to the sense in which genetics as such can be held to have an empirical warrant.

In order to discern the basic error in Eddington's critique, it is of the utmost importance to realize clearly that the thesis of the conventionality of congruence is, in the first instance, a claim concerning *structural properties of physical space and time*; only the semantical corollary of that thesis concerns the language of the geo-chronometric description of the physical world. Having failed to appreciate this fact, Eddington and those who invoke the theory of models were misled into giving a shallow caricature of the debate between the Newtonian, who affirms the factuality of congruence on the strength of the alleged intrinsicality of the metric, and his Riemannian conventionalistic critic such as Poincaré. According to the burlesqued version of this controversy, Poincaré is offering no more than a semantical truism. More specifically, the detractors suppose that their trivialization of the congruence issue can be vindicated by pointing out that we are, of course, free to decree the referents of the *unpre-empted word* "congruent," because such freedom can be exercised with respect

to any as yet semantically uncommitted term or string of symbols whatever. And, in this way, they misconstrue the conventionality of congruence as merely an inflated special case of a semantical banality holding for any and all linguistic signs or symbols, a banality which we shall call "trivial semantical conventionalism" or, in abbreviated form, "TSC."

No one, of course, will wish to deny that *qua uncommitted signs*, the terms "spatially congruent" and "temporally congruent" are fully on a par in regard to the *trivial* conventionality of the semantical rules governing their use with all linguistic symbols whatever. And thus a sensible person would hardly wish to contest that the unenlightening affirmation of the conventionality of the use of the *unpre-empted word* "congruent" is indeed a sub-thesis of TSC. But it is a serious obfuscation to identify the Riemann-Poincaré doctrine that the *ascription* of the congruence or equality *relation* to space or time intervals is conventional with the platitude that the use of the *unpre-empted word* "congruent" is conventional. And it is therefore totally incorrect to conclude that the Riemann-Poincaré tenet is merely a gratuitously emphasized special case of TSC. For what these mathematicians are advocating is not a doctrine about the semantical freedom we have in the use of the uncommitted sign "congruent." Instead, they are putting forward the initially non-semantical claim that the continua of physical space and time each lack an intrinsic metric. And the metric amorphousness of these continua then serves to *explain* that even *after* the word "congruent" has been pre-empted semantically as a spatial or temporal *equality* predicate by the axioms of congruence, congruence remains ambiguous in the sense that these axioms still allow an infinitude of mutually exclusive congruence classes of intervals. Accordingly, fundamentally *non-semantical* considerations are used to show that only a conventional choice of one of these congruence classes can provide a unique standard of length equality. In short, the conventionality of congruence is a claim not about the noise "congruent" but about the character of the conditions relevant to the *obtaining* of the kind of equality relation denoted by the term "congruent." For alternative metrization is not a matter of the freedom to use the semantically uncommitted noise "congruent" as we please; instead, it is a matter of the *non-*

uniqueness of a relation term already *pre-empted* as the *physico-spatial* (or *temporal*) equality predicate. And *this non-uniqueness arises from the lack of an intrinsic metric in the continuous manifolds of physical space and time.*

The philosophical status of the Riemann-Poincaré conventionality of congruence is fully analogous to that of Einstein's conventionality of simultaneity. And if the reasoning used by Eddington in an endeavor to establish the banality of the former were actually sound, then, as we shall now show, it would follow by a precisely similar argument that Einstein's enunciation of the conventionality of simultaneity⁹ was no more than a turgid statement of the platitude that the uncommitted word "simultaneous" (or "gleichzeitig") may be used as we please. In fact, due to the complete philosophical affinity of the conventionality of congruence with the conventionality of simultaneity, which we are about to exhibit, it will be useful subsequently to combine these two theses under the name "geo-chronometric conventionalism" or, in abbreviated form, "GC."

We saw in the case of spatial and temporal congruence that congruence is conventional in a sense other than that prior to being pre-empted semantically, the *sign* "congruent" can be used to denote anything we please. *Mutatis mutandis*, we now wish to show that precisely the same holds for the conventionality of metrical simultaneity. Once we have furnished this demonstration as well, we shall have established that neither of the component claims of conventionality in our compound GC thesis is a subthesis of TSC.

We proceed in Einstein's manner in the special theory of relativity and first decree that the *noise* "topologically simultaneous" denotes the relation of *not* being connectible by a physical causal (signal) chain, a relation which may obtain between two physical events. We now ask: is this definition unique in the sense of assuring that one and only one event at a point *Q* will be topologically simultaneous with a given event occurring at a point *P* elsewhere in space? The answer to this question depends on facts of nature, namely on the range of the causal chains existing in

⁹ A. Einstein: "On the Electrodynamics of Moving Bodies," *The Principle of Relativity, A Collection of Original Memoirs*, pp. 37-65. (New York: Dover Publications, Inc.; 1952), §1. For details, see Chapter Twelve below.

the physical world. Thus, once the above definition is given, its uniqueness is not a matter subject to regulation by semantical convention. If now we assume with Einstein as a fact of nature that light *in vacuo* is the fastest causal chain, then this postulate entails the non-uniqueness of the definition of "topological simultaneity" given above and thereby also prevents *topological* simultaneity from being a transitive relation. By contrast, if the facts of the physical world had the structure assumed by Newton, this non-uniqueness would not arise. Accordingly, the structure of the facts of the world postulated by relativity prevents the above definition of "topological simultaneity" from also serving, as it stands, as a metrical synchronization rule for clocks at the spatially separated points *P* and *Q*. Upon coupling this result with the relativistic assumption that transported clocks do not define an absolute metrical simultaneity, we see that the facts of the world leave the equality relation of metrical simultaneity *indeterminate*, for they do not confer upon topological simultaneity the uniqueness which it would require to serve as the basis of metrical simultaneity as well. Therefore, the assertion of that indeterminateness and of *the corollary that metrical simultaneity is made determinate by convention* is in no way tantamount to the purely semantical assertion that the mere uncommitted noise "metrically simultaneous" must be given a physical interpretation before it can denote, and that this interpretation is trivially a matter of convention.

Far from being a claim that a mere linguistic noise is still awaiting an assignment of semantical meaning, the assertion of the factual indeterminateness of metrical simultaneity concerns *facts of nature* which find expression in the residual non-uniqueness of the definition of "topological simultaneity" once the latter has already been given. And it is thus impossible to construe this residual non-uniqueness as being attributable to taciturnity or tight-lippedness on Einstein's part in telling us what he means by the noise "simultaneous." Here, then, we are confronted with a kind of logical gap needing to be filled by definition which is precisely analogous to the case of congruence, where the continuity of space and time issued in the residual non-uniqueness of the congruence axioms *after* they are given a spatial or temporal interpretation. When I say that metrical simultaneity is

not wholly factual but contains a conventional ingredient, what am I asserting? I am claiming none other than that the residual non-uniqueness or logical gap cannot be removed by an appeal to facts but only by a conventional choice of a unique pair of events at P and at Q as *metrically* simultaneous from within the class of pairs that are topologically simultaneous. And when I assert that it was a great philosophical (as well as physical) achievement for Einstein to have discovered the conventionality of metrical simultaneity, I am crediting Einstein *not* with the triviality of having decreed semantically the meaning of the *noise* "metrically simultaneous" (or "gleichzeitig") but with the recognition that, contrary to earlier belief, the facts of nature are such as to deny the required kind of semantical univocity to the already pre-empted term "simultaneous" ("gleichzeitig"). In short, Einstein's insight that metrical simultaneity is conventional is a contribution to the theory of time rather than to semantics, because *it concerns the character of the conditions relevant to the obtaining of the relation denoted by the term "metrically simultaneous."*

The conventionality of metrical simultaneity has just been formulated without any reference whatever to the relative motion of different Galilean frames, and does not depend upon there being a relativity or non-concordance of simultaneity as between different Galilean frames. On the contrary, as Chapter Twelve will show in detail, it is the conventionality of simultaneity which provides the logical framework within which the relativity of simultaneity can first be understood: if each Galilean observer adopts the particular metrical synchronization rule adopted by Einstein in Section 1 of his fundamental paper¹—a rule which corresponds to the value $\epsilon = \frac{1}{2}$ in the Reichenbach notation²—then the relative motion of Galilean frames issues in their choosing as metrically simultaneous *different pairs of events* from within the class of topologically simultaneous events at P and Q , a result embodied in the familiar Minkowski diagram.

In discussing the definition of simultaneity,³ Einstein italicized

¹ *Ibid.*

² H. Reichenbach: *The Philosophy of Space and Time* (New York: Dover Publications, Inc.; 1957), p. 127.

³ A. Einstein: "On the Electrodynamics of Moving Bodies," *The Principle of Relativity, A Collection of Original Memoirs, op. cit.*, §1.

the words "by definition" in saying that the *equality* of the to and fro velocities of light between two points *A* and *B* is a matter of definition. Thus, he is asserting that metrical simultaneity is a matter of definition or convention. Do the detractors really expect anyone to believe that Einstein put these words in italics to convey to the public that the *noise* "simultaneous" can be used as we please? Presumably they would recoil from this conclusion. But how else could they solve the problem of making Einstein's avowedly conventionalist conception of metrical simultaneity compatible with their semantical trivialization of GCC? H. Putnam, one of the advocates of the view that the conventionality of congruence is a subthesis of TSC, has sought to meet this difficulty along the following lines: in the case of the congruence of intervals, one would never run into trouble upon using the *customary* definition;⁴ but in the case of simultaneity, actual contradictions would be encountered upon using the customary classical definition of metrical simultaneity, which is based on the transport of clocks and is vitiated by the dependence of the clock rates (readings) on the transport velocity.

But Putnam's retort will not do. For the appeal to Einstein's recognition of the inconsistency of the classical definition of metrical simultaneity accounts only for his abandonment of the latter but does not illuminate—as does the thesis of the conventionality of simultaneity—the logical status of the *particular set* of definitions which Einstein put in its place. Thus, the Putnamian retort does not recognize that the logical status of Einstein's synchronization rules is not at all adequately rendered by saying that whereas the classical definition of metrical simultaneity was inconsistent, Einstein's rules have the virtue of consistency. For what needs to be elucidated is *the nature of the logical step* leading to Einstein's particular synchronization scheme within the wider framework of the *alternative consistent sets of rules* for metrical simultaneity any one of which is allowed by the non-uniqueness of topological simultaneity. Precisely this elucidation is given, as we have seen, by the thesis of the conventionality of metrical simultaneity.

We see therefore that the philosophically illuminating con-

⁴ We shall see in our discussion of time measurement on the rotating disk of the general theory of relativity in Chapter Two that there is one sense of "trouble" for which Putnam's statement would not hold.

ventionality of an affirmation of the congruence of two intervals or of the metrical simultaneity of two physical events does not inhere in the arbitrariness of what linguistic sentence is used to express the proposition that a relation of equality obtains among intervals, or that a relation of metrical simultaneity obtains between two physical events. Instead, the important conventionality lies in the fact that even *after* we have specified what respective linguistic sentences will express these propositions, a convention is ingredient in each of the propositions expressed, i.e., in the very obtaining of a congruence relation among intervals or of a metrical simultaneity relation among events.

These considerations enable us to articulate the misunderstanding of the conventionality of congruence to which Eddington and contemporary proponents of its *model-theoretic* trivialization fell prey. It will be recalled that the latter critics argued somewhat as follows: "The theory of models of uninterpreted formal calculi shows that there can be no basis at all for an epistemological distinction *within* the system of physical geometry (or chronometry) between factual statements, on the one hand, and supposedly conventional statements of rigidity (or isochronism), on the other. For we are just as free to give a *non*-customary spatial interpretation of, say, the abstract sign 'point' in the formal geometrical calculus as of the sign 'congruent,' and hence the physical interpretation of the relation term 'congruent' (for line segments) cannot occupy an epistemologically distinguished position among the semantical rules effecting the interpretation of the formal system, all of which are on a par in regard to conventionality."

But this objection overlooks that (a) the obtaining of the spatial congruence relation provides scope for the role of convention because, independently of the particular formal geometrical calculus which is being interpreted, the term "congruent" functions as a spatial *equality predicate* in its *non*-customary spatial interpretations no less than in its customary ones; (b) consequently, suitable alternative spatial interpretations of the term "congruent" and correlatively of "straight line" ("geodesic") show that it is always a live option (subject to the restrictions imposed by the existing topology) to give either a Euclidean or a *non*-Euclidean description of the same body of physico-geo-

metrical facts; and (c) by contrast, the possibility of alternative spatial interpretations of such other primitives of rival geometrical calculi as "point" does not generally issue in this option. Our concern is to note that, even disregarding inductive imprecision, the empirical facts themselves do not uniquely dictate the truth of either Euclidean geometry or of one of its *non-Euclidean* rivals in virtue of the lack of an intrinsic metric. Hence in this context the different spatial interpretations of the term "congruent" (and hence of "straight line") in the respective geometrical calculi play a philosophically different role than do the interpretations of such other primitives of these calculi as "point," since the latter generally have the same spatial meaning in both the Euclidean and non-Euclidean descriptions. The pre-eminent status occupied by the interpretation of "congruent" in this context becomes apparent once we cease to look at physical geometry as a spatially interpreted system of abstract *synthetic* geometry and regard it instead as an interpreted system of abstract *differential* geometry of the Gauss-Riemann type: by choosing a particular distance function $ds = \sqrt{g_{ik}dx^i dx^k}$ for the line element, we specify not only what segments are congruent and what lines are straights (geodesics) but the entire geometry, since the metric tensor g_{ik} fully determines the Gaussian curvature K . To be sure, if one were discussing not the alternative between a Euclidean and a non-Euclidean description of the same spatial facts but rather the set of all models (including non-spatial ones) of a given calculus, say the Euclidean one, then indeed the physical interpretation of "congruent" and of "straight line" would not merit any more attention than that of other primitives like "point."⁵

II. Putnam and P. K. Feyerabend have elaborated the following corollary of Eddington's charge of triviality: GC must be a

⁵ We have been speaking of certain *uninterpreted* formal calculi as systems of synthetic or differential *geometry*. It must be understood, however, that *prior* to being given a *spatial* interpretation, these abstract deductive systems no more qualify as *geometries*, strictly speaking, than as systems of genetics or of anything else; they are called "geometries," it would seem, only "because the name seems good, on emotional and traditional grounds, to a sufficient number of competent people." (O. Veblen and J. H. C. Whitehead: *The Foundations of Differential Geometry*, No. 29 of "Cambridge Tracts in Mathematics and Mathematical Physics" [Cambridge: Cambridge University Press; 1932], p. 17.)

subthesis of TSC because GC has bona fide analogues in every branch of human inquiry, such that GC cannot be construed as an insight into the structure of space or time. As Eddington puts it:

The law of Boyle states that the pressure of a gas is proportional to its density. It is found by experiment that this law is only approximately true. A certain mathematical simplicity would be gained by conventionally redefining *pressure* in such a way that Boyle's law would be rigorously obeyed. But it would be high-handed to appropriate the word *pressure* in this way, unless it had been ascertained that the physicist had no further use for it in its original meaning.⁹

P. K. Feyerabend has noted that what Eddington seems to have in mind here is the following: instead of revising Boyle's law

$$pv = RT$$

in favor of van der Waals's law

$$(p + \frac{a}{v^2})(v - b) = RT,$$

we could preserve the statement of Boyle's law by merely redefining "pressure"—now to be symbolized by "P" in its new usage—putting

$$P =_{\text{def}} (p + \frac{a}{v^2}) (1 - \frac{b}{v}).$$

In the same vein, H. Putnam maintains that instead of using phenomenalist (naive realist) color words as we do customarily in English, we could adopt a new usage for such words—to be called the "Spenglish" usage—as follows: we take a white piece of chalk, for example, which is moved about in a room, and we lay down the rule that depending (in some specified way) upon the part of the visual field which its appearance occupies, its color will be called "green," "blue," "yellow," etc. rather than "white" under constant conditions of illumination.

It is a fact, of course, that whereas actual scientific practice in the general theory of relativity, for example, countenances and uses remetrizational procedures based on non-customary

⁹ A. S. Eddington, *Space, Time, and Gravitation*, *op. cit.*, p. 10.

definitions of temporal congruence;⁷ scientific practice does not contain any examples of Putnam's "Spenglish" *space-dependent* (or *time-dependent*) use of phenomenalist (naïve realist) *color predicates* to denote the color of a given object in various places (or at various times) under like conditions of illumination. According to Eddington and Putnam, the existence of non-customary usages of "congruent" in the face of there being no such usages of color predicates is no more than a fact about the linguistic behavior of the members of our linguistic community. We saw that the use of linguistic alternatives in the specifically geo-chronometric contexts reflects fundamental *structural properties* of the facts to which these alternative descriptions pertain. And we must now demonstrate that Eddington's and Putnam's alleged analogues of GC are *pseudo-analogues* in the sense of failing to show that every empirical domain possesses analogues to GC. I should emphasize, however, that my rejection of the aforementioned purported analogues is not intended to deny the existence of one or another genuine analogue but to deny only that GC may be deemed to be trivial on the strength of such relatively few bona fide analogues as may obtain.

The essential point in assessing the cogency of the purported analogues is the following: do the domains from which they are drawn (e.g., phenomenalist or naïve realist color properties, pressure phenomena, etc.) exhibit *structural* counterparts to (a) those factual properties of the world postulated by relativity which entail the non-uniqueness of topological simultaneity, and (b) the postulated topological properties of physical space and time which make for the non-uniqueness of the respective spatial and temporal interpretations allowed by the congruence axioms? Or are the examples cited by Eddington and Putnam analogues of the conventionality of metrical simultaneity or of congruence

⁷ The proponents of ordinary language usage in science, to whom the "ordinary man" seems to be the measure of all things, may wish to rule out *non-customary* congruence definitions as linguistically illegitimate. But they would do well to remember that it is no more incumbent upon the scientist (or philosopher of science) to use the customary scientific definition of congruence in *every* geo-chronometric description than it is obligatory for, say, the student of mechanics to be bound by the familiar common sense meaning of "work," which contradicts the mechanical meaning as given by the space integral of the force.

only in the impoverished, trivial sense that they feature linguistically alternative equivalent descriptions while lacking the following decisive property of the geo-chronometric cases: the alternative metrizations are the linguistic renditions or reverberations, as it were, of the *structural properties* assuming the aforementioned two kinds of non-uniqueness enunciated by GC? If the examples given are analogues only in the superficial, impoverished sense—as indeed I shall show them to be—then what have Eddington and Putnam accomplished by their examples? In that case they have merely provided unnecessary illustrations of the correctness of TSC without proving their examples to be on a par with the geo-chronometric ones. In short, their examples will then have served in no way to make good their claim that GC is a subthesis of TSC.

We shall find presently that their examples fail because (a) the domains to which they pertain do not exhibit structural counterparts to those features of the world which make the definitions of topological simultaneity and the axioms for spatial or temporal congruence *non-unique*, and (b) Putnam's example in "Spenglish" is indeed an illustration only of the trivial conventionality of all language: no structural property of the domain of phenomenal color (e.g., in the appearances of chalk) is rendered by the feasibility of the Spenglish description.

To state my objections to the Eddington-Putnam thesis, I call attention to the following two sentences:

(A) Person X does not have a gall bladder.

(B) The platinum-iridium bar in the custody of the Bureau of Weights and Measures in Paris (Sèvres) is one meter long everywhere rather than some other number of meters (after allowance for "differential forces").

I maintain that there is a *fundamental difference* between the senses in which each of these statements can possibly be held to be conventional, and I shall refer to these respective senses as "A-conventional" and "B-conventional": in the case of statement (A), what is conventional is only the use of the given sentence to render the proposition of X's not having a gall bladder, not the factual proposition expressed by the sentence. This A-conventionality is of the trivial weak kind affirmed by TSC. On the other hand, (B) is conventional not merely in the trivial sense

that the English sentence used could have been replaced by one in French or some other language but in the much stronger and deeper sense that it is not a factual proposition that the Paris bar has everywhere a length unity in the meter scale even after we have specified what sentence or string of noises will express this proposition. In brief, in (A), semantic conventions are *used*, whereas, in (B), a semantic convention is *mentioned*. Now I claim that the alleged analogues of Eddington and Putnam illustrate conventionality only in the sense of A-conventionality and therefore cannot score against my contention that geo-chronometric conventionality is *non-trivial* by having the character of B-conventionality.

Specifically, I assert that statements about phenomenalist colors are empirical statements pure and simple in the sense of being only A-conventional and not B-conventional, while an important class of statements of geo-chronometry possess a different, deeper conventionality of their own by being B-conventional. What is it that is conventional in the case of the color of a given piece of chalk, which appears white in various parts of the visual field? I answer: only our customary decision to use the same word to refer to the various qualitatively same white chalk appearances in different parts of the visual field. But it is not conventional whether the various chalk appearances do have the same phenomenal color property (to within the precision allowed by vagueness) and thus are "color-congruent" to one another or not! Only the color-words are conventional, not the obtaining of specified color-properties and of color-congruence. And the obtaining of color-congruence is non-conventional quite independently of whether the various occurrences of a particular shade of color are denoted by the same color-word or not. In other words, there is no convention in whether two objects or two appearances of the same object under like optical conditions have the same phenomenal color property of whiteness (apart from vagueness) but only in whether the noise "white" is applied to both of these objects or appearances, to one of them and not to the other (as in Putnam's chalk example) or to neither. And the alternative color descriptions *do not render any structural facts of the color domain* and are therefore purely trivial.

Though failing in this decisive way, Putnam's chalk color case

is falsely given the *semblance* of being a bona fide analogue to the spatial congruence case by the device of laying down a rule which makes the use of color names *space-dependent*: the rule is that different noises (color names) will be used to refer to the same de facto color property occurring in different portions of visual space. But this stratagem cannot overcome the fact that while the assertion of the possibility of assigning a space-dependent length to a transported rod reflects linguistically the objective non-existence of an intrinsic metric, the space-dependent use of color names does not reflect a corresponding property of the domain of phenomenal colors in visual space. In short, the phenomenalist color of an appearance is an intrinsic, objective property of that appearance, and phenomenal color-congruence is an objective relation (to within the precision allowed by vagueness). But the length of a body and the congruence of non-coinciding intervals are not similarly non-conventional. And we shall see in our critique of Whitehead that this conclusion cannot be invalidated by the fact that two non-coinciding intervals can *look* spatially-congruent no less than two color patches can appear color-congruent.

Next consider Eddington's example of the preservation of the language of Boyle's law to render the *new facts* affirmed by van der Waals's law by the device of giving a new meaning to the word "pressure" as explained earlier. The customary concept of pressure has geo-chronometric ingredients (force, area), and any alterations made in the geo-chronometric congruence definitions will, of course, issue in changes as to what pressures will be held to be equal. But the conventionality of the geo-chronometric ingredients is not of course at issue, and we ask: of what *structural feature* of the domain of pressure phenomena does the possibility of Eddington's above linguistic transcription render testimony? The answer clearly is: *of none*. Unlike GC, the thesis of the "conventionality of pressure," if put forward on the basis of Eddington's example, concerns only A-conventionality and is thus merely a special case of TSC. We observe incidentally, that two pressures which are equal on the customary definition will also be equal (congruent) on the suggested redefinition of that term: apart from the distinctly geo-chronometric ingredients not here at issue, the domain of pressure phenomena does not pre-

sent us with any structural property as the counterpart of the lack of an intrinsic metric of space which would be reflected by the alternative definitions of "pressure."

The absurdity of likening the conventionality of spatial or temporal congruence to the conventionality of the choice between the two above meanings of "pressure" or between English and Spenglish color discourse becomes patent upon considering the expression given to the conventionality of congruence by the Klein-Lie group-theoretical treatment of congruences and metric geometries. For their investigations likewise serve to show, as we shall now indicate, how far removed from being a semantically uncommitted noise the term "congruent" is while still failing to single out a unique congruence class of intervals, and how badly amiss it is for Eddington and Putnam to maintain that this non-uniqueness is merely a special example of the semantical non-uniqueness of all uncommitted noises.

Felix Klein's Erlangen Program (1872) of treating geometries from the point of view of groups of spatial transformations was rooted in the following two observations: first, the properties in virtue of which spatial congruence has the logical status of an equality relation depend upon the fact that displacements are given by a group of transformations, and second, the congruence of two figures consists in their being intertransformable into one another by means of a certain transformation of points. Continuing Klein's reasoning, Sophus Lie then showed that, in the context of this group-theoretical characterization of metric geometry, the conventionality of congruence issues in the following results: first, the set of all the continuous groups in space having the property of displacements in a bounded region fall into three types which respectively characterize the geometries of Euclid, Lobachevski-Bolyai, and Riemann,⁸ and second, for each of these metrical geometries, there is not one but an infinitude of different congruence classes,⁹ a result which we shall demonstrate in Chapter Three without group-theoretical devices. On the Eddington-Putnam thesis, Lie's profound and justly celebrated results

⁸ R. Bonola: *Non-Euclidean Geometry* (New York: Dover Publications, Inc.; 1955), p. 153.

⁹ Cf. A. N. Whitehead: *The Principle of Relativity*, *op. cit.*, Chapter iii, p. 49.

no less than the relativity of simultaneity and the conventionality of temporal congruence must be consigned absurdly to the limbo of trivial semantical conventionality along with Spenglish color discourse!

As previously noted, these objections against the Eddington-Putnam claim that GC has bona fide analogues in every empirical domain are not intended to deny the existence of one or another genuine analogue but to deny only that GC may be deemed to be trivial on the strength of such relatively few bona fide analogues as may obtain. Putnam has given one example which does seem to qualify as a bona fide analogue. This example differs from his color case in that not merely the name given to a property but the sameness of the property named is dependent on spatial position as follows: when two bodies are at essentially the same place, their sameness with respect to a certain property is a matter of fact, but when they are (sufficiently) apart spatially, no objective relation of sameness or difference with respect to the given property obtains between them. And, in the latter case, therefore, it becomes a matter of convention whether sameness or difference is ascribed to them in this respect.

Specifically, suppose that we do not aim at a definition of mass adequate to classical mechanics and thus ignore Mach's definition of mass.¹ Then we can consider Putnam's hypothetical definition of "mass-equality," according to which two bodies balancing one another on a suitable scale at what is essentially the same place in space have equal masses. Whereas on the Machian definition mass equality obtains between two bodies *as a matter of fact* independently of the extent of their spatial separation, on Putnam's definition such separation leaves the relation of mass equality at a distance *indeterminate*. Hence, on Putnam's definition, it would be a matter of convention whether (a) we would say that two masses which balance at a given place remain equal to one another in respect to mass after being spatially separated, or (b) we would make the masses of two bodies space-dependent such that two masses that balance at one place would have different masses when separated, as specified by a certain function of the coordinates. The conventionality arising

¹ For a concise account of that definition, see L. Page: *Introduction to Theoretical Physics*, *op. cit.*, pp. 56-58.

in Putnam's mass example is not a consequence of GC but is logically independent of it. For it is not spatial congruence of non-coinciding intervals but spatial position that is the source of conventionality here.

The bona fide character of Putnam's mass analogue cannot, however, invalidate our earlier conclusion that we must attach a very different significance to alternative metric geometries or chronometries as equivalent descriptions of the same facts than to alternate types of visual color discourse as equivalent descriptions of the same phenomenal data. By the same token, we must attach much greater significance to being able to render factually different geo-chronometric states of affairs by the same geometry or chronometry, coupled with appropriately different congruence definitions, than to formulating both Boyle's law and van der Waals's law, which differ in factual content, by the same law-statement coupled with appropriately different semantical rules. In short, there is an important respect in which physical geo-chronometry is less empirical than all or almost all of the non-geo-chronometric portions (ingredients) of other sciences.

(E) BRIDGMAN

We have grounded the conventionality of spatial and temporal congruence on the continuity of the manifolds of space and time. And, in thus arguing that "true," absolute, or intrinsic rigidity and isochronism are non-existing properties in these respective continua, we did not adduce any homocentric-operationist criterion of factual meaning. For we did not say that the actual and potential failure of human testing operations to disclose "true rigidity" constitutes either its non-existence or its meaninglessness. It will be well, therefore, to compare our Riemannian espousal of the conventionality of rigidity and isochronism with the reasoning of those who arrive at this conception of congruence by arguing from non-testability or from an operationist view of scientific concepts. Thus W. K. Clifford writes: "we have defined length or distance by means of a measure which can be carried about *without changing its length*. But how then is this property of the measure to be tested? . . . Is it possible . . . that lengths do really change by mere moving

about, without our knowing it? Whoever likes to meditate seriously upon this question will find that it is wholly devoid of meaning."²

We saw that within our Riemannian framework of ideas, length is *relational* rather than absolute in a twofold sense: First, length obviously depends numerically on the units used and is thus arbitrary to within a constant factor, and second, in virtue of the lack of an intrinsic metric, sameness or change of the length possessed by a body in different places and at different times consist in the sameness or change respectively of the ratio (relation) of that body to the conventional standard of congruence. Whether or not this ratio changes is quite independent of any human discovery of it: the number of times which a given body *B* will contain a certain unit rod is a property of *B* that is not first conferred on *B* by human operations. As Reichenbach has noted: "*The objective character of the physical statement [concerning the geometry of physical space] is thus shifted to a statement about relations . . . it is a statement about a relation between the universe and rigid rods.*"³ And thus the relational character of length derives, in the first instance, not from how we human beings measure lengths but from the failure of the continuum of physical space to possess an intrinsic metric, a failure obtaining quite independently of our measuring activities. In fact, it is this relational character of length which prescribes and regulates the kinds of human operations appropriate to its discovery. Since, to begin with, there exists no property of true rigidity to be disclosed by any human test, no test could possibly reveal its presence. Accordingly, the unascertainability of true rigidity by us humans is a consequence of its non-existence in physical space and evidence for that non-existence but not constitutive of it.

On the basis of this non-homocentric relational conception of length, the utter vacuousness of the following assertion is evident at once: overnight *everything* has *expanded* (i.e., increased its length) but such that all length *ratios* remained unaltered. That

² W. K. Clifford: *The Common Sense of the Exact Sciences* (New York: Dover Publications, Inc.; 1955), pp. 49–50.

³ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 37; italics are in the original.

such an alleged "expansion" will elude any and all human tests is then obviously explained by its not having obtained: the increase in the ratios between all bodies and the congruence standard which would have constituted the expansion avowedly did not materialize. And in the absence of an intrinsic metric, it is sheer nonsense to say that overnight the Paris meter ceased to be truly one meter and is now only apparently so, since it has actually expanded.

We see that the relational theory of length and hence the particular assertion of the vacuousness of a universal nocturnal expansion do not depend on a grounding of the meaning of the metrical concepts of length and duration on human testability or manipulations of rods and clocks in the manner of Bridgman's homocentric operationism. Moreover, there is a further sense in which the Riemannian recognition of the need for a specification of the congruence criterion does not entail an operationist conception of congruence and length. As we noted preliminarily at the beginning of this chapter and will see in detail in Chapter Four, the definition of "congruence" on the basis of the coincidence behavior common to all kinds of transported solid rods provides a rule of correspondence (coordinative definition) *through the mediation of hypotheses and laws* that are *collateral* to the abstract geometry receiving a physical interpretation. For the physical laws used to compute the corrections for thermal and other substance-specific deformations of solid rods made of different kinds of materials enter integrally into the definition of "congruent." Thus, in the case of "length" no less than in most other cases, operational definitions (in any distinctive sense of the term "operational") are a quite idealized and limiting species of correspondence rule even though the definition of "length" is often adduced as the prototype of all "operational" definitions in Bridgman's sense.

Further illustrations of this fact are given by Reichenbach, who cites the definitions of the unit of length on the basis of the wave length of cadmium light and also in terms of a certain fraction of the circumference of the earth and writes: "Which distance serves as a unit for actual measurements can ultimately be given only by reference to some actual distance. . . . We say with regard to the measuring rod . . . that only 'ultimately' the

reference is to be conceived in this form, because we know that by means of the interposition of conceptual relations the reference may be rather remote."⁴ An even stronger repudiation of the operationist account of the definition of "length" because of its failure to allow for the role of auxiliary theory is presented by K. R. Popper, who says: "As to the doctrine of operationalism—which demands that scientific terms, such as length . . . should be defined in terms of the appropriate experimental procedure—it can be shown quite easily that all so-called operational definitions will be circular . . . the circularity of the operational definition of length . . . may be seen from the following facts: (a) the 'operational' definition of *length* involves *temperature* corrections, and (b) the (usual) operational definition of *temperature* involves measurements of length."⁵

See Append. §2

(F) RUSSELL

During the years 1897–1900, B. Russell and H. Poincaré had a controversy which was initiated by Poincaré's review⁶ of Russell's *Foundations of Geometry* of 1897, and pursued in Russell's reply⁷ and Poincaré's rejoinder.⁸ Russell criticized Poincaré's conventionalist conception of congruence and invoked the existence of an intrinsic metric as follows:

It seems to be believed that since measurement [i.e., comparison by means of the congruence standard] is necessary to *discover* equality or inequality, these cannot exist without measurement. Now the proper conclusion is exactly the opposite. Whatever one can discover by means of an operation must exist independently of that operation: America existed before Chris-

⁴ *Ibid.*, p. 128.

⁵ K. R. Popper: *The Logic of Scientific Discovery* (London: Hutchinson and Co., Ltd.; 1959), pp. 440 and 440n. In Chapter Four, we shall see how the circularity besetting the operationist conception of length is handled within our framework when making allowance for thermal and other deformations in the statement of the definition of congruence.

⁶ H. Poincaré: "Des Fondements de la Géométrie, à propos d'un Livre de M. Russell," *Revue de Métaphysique et de Morale*, Vol. VII (1899), pp. 251–79.

⁷ B. Russell: "Sur les Axiomes de la Géométrie," *Revue de Métaphysique et de Morale*, Vol. VII (1899), pp. 684–707.

⁸ H. Poincaré: "Sur les Principes de la Géométrie, Réponse à M. Russell," *Revue de Métaphysique et de Morale*, Vol. VIII (1900), pp. 73–86.

topher Columbus, and two quantities of the same kind must *be* equal or unequal before being measured. Any method of measurement [i.e., any congruence definition] is good or bad according as it yields a result which is true or false. Mr. Poincaré, on the other hand, holds that measurement creates equality and inequality. It follows [then] . . . that there is nothing left to measure and that equality and inequality are terms devoid of meaning.⁹

The congruence issue between Russell and Poincaré here is, of course, distinct from the controversy between the neo-Kantian and the empiricist conceptions of geometry, which plays a major role in Russell's *Foundations of Geometry* of 1897.

As we recall from our citation of Riemann's "Inaugural Lecture" in Section B, if space were discrete in some specified sense, then the "distance" between two elements could be defined intrinsically in a rather natural way by the cardinality of the least number of intervening elements. In that eventuality, both the "length" of any given interval and the obtaining of spatial congruence among separated intervals would be wholly independent of the behavior of any transported standard. And if space is thus granular, the logic of the discovery of length would be analogous to that of Columbus's discovery of America in Russell's example, the role of the measuring rod then being no more than a purely epistemic one. Moreover, in the case of discreteness, the measuring rod is even dispensable for epistemic purposes since a separate determination of the number of chunks or space-atoms contained in each of two bodies would yield a verdict on their spatial congruence before any comparison of them would need to be effected via a transported congruence standard. Russell overlooked that once we assume the continuity of our physical space, the congruence of two line segments cannot derive from their respective possession of an intrinsic metric attribute and that their congruence depends for its very obtaining and not merely for its human ascertainment on a relation to an *extrinsic* standard whose "rigidity" under transport is decreed conventionally. It is the bodies or segments themselves, but not their relations of spatial equality or inequality which exist independently of the coincidence behavior of a

⁹ B. Russell: "Sur les Axiomes de la Géométrie," *op. cit.*, pp. 687-88.

transported congruence standard. And *to recognize the dependence of the very obtaining of spatial congruence among separated intervals on the mediation of the transported congruence standard is not to confuse measurement in its epistemic function of discovery with the facts ascertained by it.* Hence we see that Poincaré was not guilty of the following error: what makes the property of length in our supposedly continuous physical space different from those discovered by Columbus in Russell's example is first generated by the difference in the respective operational procedures used by us humans in their discovery. Instead, Poincaré rested his case on the pre-existing difference in the properties to be discovered, a difference that determines and lends significance to the operational procedures appropriate to their discovery.

Although we therefore reject Russell's argument against Poincaré, our critique of the model-theoretic trivialization of the conventionality of congruence shows that we must likewise reject as inadequate the following kind of criticism of Russell's position, which he would have rightly regarded as a *petitio principii*: "Russell's claim is an absurdity, because it is the *denial* of the truism that we are at liberty to give whatever physical interpretations we like to such abstract signs as 'congruent line segments' and 'straight line' and then to inquire whether the system of objects and relations thus arbitrarily named is a model of one or another abstract geometric axiom system. Hence, purely linguistic considerations suffice to show that there can be no question, as Russell would have it, whether two non-coinciding segments are truly equal or not and whether measurement is being carried out with a standard yielding results that are true in that sense. Accordingly, awareness of the model-theoretic conception of geometry would have shown Russell that (1) alternative metrization of spatial and temporal continua should never have been either startling or a matter for dispute, (2) the stake in his controversy with Poincaré was no more than the pathetic one that Russell was advocating the customary linguistic usage of the term "congruent" (for line segments) while Poincaré was maintaining that we need not be bound by the customary usage but are at liberty to introduce bizarre ones as well."

Since this model-theoretic argument altogether fails to come

to grips with Russell's root-assumption of an intrinsic metric, he would have been entitled to dismiss it as a shallow *petitio* by raising exactly the same objections against the alternative metrizable-ness of space and time, which we attributed to the Newtonian at the end of Section C above. And Russell might have gone on to point out that the model-theoretician cannot evade the spatial equality issue by (i) noting that there are axiomatizations of each of the various geometries dispensing with the abstract relation term "congruent" (for line segments), and (ii) claiming that there can then be no problem as to what physical interpretations of that relation term are permissible. For a metric geometry makes metrical comparisons of equality and inequality, however covertly or circuitously these may be rendered by its language. It is quite immaterial, therefore, whether the relation of spatial equality between line segments is designated by the term "congruent" or by some other term or terms. Thus, for example, Tarski's axioms for elementary Euclidean geometry¹ do not employ the term "congruent" for this purpose, using instead a quaternary predicate denoting the equidistance relation between four points. Also, in Sophus Lie's group-theoretical treatment of metric geometries, the congruences are specified by groups of point transformations.² But just as Russell invoked his conception of an intrinsic metric to restrict the permissible spatial interpretations of "congruent line segments," so also he would have maintained that it is never arbitrary what quartets of physical points may be regarded as the denotata of Tarski's quaternary equidistance predicate. And he would have imposed corresponding restrictions on Lie's transformations, since the displacements defined by these groups of transformations have the logical character of spatial congruences. These considerations show that it will not suffice in this context simply to take the model-theoretic conception of geometry for granted and thereby to dismiss Russell's claim peremptorily in favor of alternative metrizable-ness. Rather what is needed is a *refutation of Russell's root-assumption of an intrinsic metric*: to exhibit the untenability

¹ A. Tarski: "What is Elementary Geometry?" *The Axiomatic Method*, ed. by L. Henkin, P. Suppes, and A. Tarski (Amsterdam: North Holland Publishing Company; 1959), pp. 16–29.

² R. Bonola: *Non-Euclidean Geometry*, *op. cit.*, pp. 153–54.

of that assumption as we have endeavored to do earlier is to provide the physical justification of the model-theoretic affirmation that a given set of physico-spatial facts of coincidence of a transported rod may be held to be as much a realization of a Euclidean calculus as of a non-Euclidean one yielding the same topology.

(C) WHITEHEAD

A. N. Whitehead has given a *perceptualistic* version of Russell's argument by attempting to ground an intrinsic metric of physical space and time on the deliverances of sense. We therefore now turn to an examination of Whitehead's conception of congruence.

Commenting on the Russell-Poincaré controversy, Whitehead³ maintains the following: First, Poincaré's argument on behalf of alternative metrizableability is unanswerable only if the philosophy of physical geometry and chronometry is part of an epistemological framework resting on an illegitimate bifurcation of nature; second, consonant with the rejection of bifurcation, we must ground our metric account of the space and time of nature not on the relations between material bodies and events as fundamental entities but on the more ultimate metric deliverances of sense perception; and third, perceptual time and space exhibit an intrinsic metric. Specifically, by countenancing the bifurcation of nature to begin with, Riemann was driven to the conclusion that the very meaning of spatial (or temporal) congruence depends on a standard which cannot make any claim to uniqueness. On his bifurcationist view that the congruence standard must "come from somewhere else," the congruence thus defined can enjoy only a conventional pre-eminence over alternative congruences that might have been selected with equal mathematical justification. Contrariwise, an antibifurcationist theory of nature relies on the immediate deliverances of sense awareness which furnish and justify a unique set of physically relevant congruence relations for both space and time (to within the precision allowed by sensory vagueness). And by thus taking cognizance of the disclosures of sense perception, we can confer

³ A. N. Whitehead: *The Concept of Nature* (Cambridge: Cambridge University Press; 1926), pp. 121-24. Hereafter this work will be cited as *CN*.

intelligibility on the unique role of the familiar congruence criteria for space and time in the face of the mathematically competing claims of "the indefinite herd" of other, mutually exclusive congruences.⁴

Clarity may be served by first stating in turn the gist of the several arguments given by Whitehead in support of his contentions. After thus obtaining a synoptic view of his polemic, we shall cite and examine each of his arguments in some detail.

According to Whitehead, the thesis of alternative metrizableability can be seen to be absurd for two reasons: First, the conventionalist conception defines temporal congruence by the requirement that Newton's laws (as modified by the small corrections for the relativistic motion of the perihelia) are true. But "uniformity in change is directly perceived"⁵ and "the measurement of time was known to all civilized nations long before the laws [of Newton] were thought of. It is this time as thus measured that the laws are concerned with. . . . It is for science to give an intellectual account of what is so evident in sense awareness."⁶ Second, just as it is an objective datum of experience that two phenomenal color patches have the same color, i.e., are "color-congruent" to within the precision allowed by vagueness, so also we *see* that a given rod has the same length in different positions, thereby showing that spatial congruence or *matching* is no less objective a relation than phenomenal color-"congruence."⁷ Hence, "it is a fact of nature that a distance of thirty miles is a long walk for anyone. There is no convention about that."⁸ And it is "no slight recommendation"⁹ of the antibifurcationist theory of nature that it removes the following difficulty besetting the bifurcationist version of classical nineteenth-century geometry and physics: There is a "breakdown of the uniqueness of congruence for space [because of alternative metrizableability] and of its very existence for time," because "time, in itself, according

⁴ *Ibid.*, p. 124.

⁵ *Ibid.*, p. 137.

⁶ *Ibid.*, p. 140.

⁷ Cf. A. N. Whitehead: *The Principles of Natural Knowledge* (Cambridge: Cambridge University Press; 1955), p. 56.

⁸ Cited from Whitehead in R. M. Palter: *Whitehead's Philosophy of Science* (Chicago: University of Chicago Press; 1960), p. 93.

⁹ A. N. Whitehead: *CN*, *op. cit.*, p. 124.

to the classical theory, presents us with no qualifying [i.e. congruence] class at all.¹ Furthermore, contrary to the "modern doctrine that 'congruence' means the possibility of coincidence," the correct account of the matter is that "although 'coincidence' is used as a *test* of congruence, it is not the *meaning* of congruence."² Also "immediate judgments of congruence are presupposed in measurement, and the process of measurement is merely a procedure to extend the recognition of congruence to cases where these immediate judgments are not available. Thus we cannot define congruence by measurement."³

Turning now to the detailed examination of Whitehead's argumentation, we note that he proposes to point out "the factor in nature which issues in the pre-eminence of one [spatial] congruence relation over the indefinite herd of other such relations"⁴ as follows:

The reason for this result is that nature is no longer confined within space at an instant. Space and time are now interconnected; and this peculiar factor of time which is so immediately distinguished among the deliverances of our sense awareness, relates itself to one particular congruence relation in space.⁵ . . . Congruence depends on motion, and thereby is generated the connexion between spatial congruence and temporal congruence.⁶

Whitehead's argument is thus seen to turn on his ability to show that *temporal* congruence cannot be regarded as conventional in physics as understood by Riemann and Poincaré. He believes to have justified this crucial claim by the following reasoning in which he refers to the conventionalist conception as "the prevalent view" and to his opposing thesis as "the new theory":

The new theory provides a definition of the congruence of periods of time. The prevalent view provides no such definition.

¹ A. N. Whitehead: *The Principle of Relativity* (Cambridge: Cambridge University Press; 1922), p. 49. Hereafter this work will be cited as *R*.

² A. N. Whitehead: *Process and Reality* (New York: The Macmillan Co.; 1929), p. 501. Hereafter this work will be cited as *PR*.

³ A. N. Whitehead: *CN, op. cit.*, p. 121.

⁴ *Ibid.*, p. 124.

⁵ *Ibid.*

⁶ *Ibid.*, p. 126.

Its position is that if we take such time-measurements so that certain familiar velocities which seem to us to be uniform are uniform, then the laws of motion are true. Now in the first place no change could appear either as uniform or non-uniform without involving a definite determination of the congruence for time-periods. So in appealing to familiar phenomena it allows that there is some factor in nature which we can intellectually construct as a congruence theory. It does not however say anything about it except that the laws of motion are then true. Suppose that with some expositors we cut out the reference to familiar velocities such as the rate of rotation of the earth. We are then driven to admit that there is no meaning in temporal congruence except that certain assumptions make the laws of motion true. Such a statement is historically false. King Alfred the Great was ignorant of the laws of motion, but knew very well what he meant by the measurement of time, and achieved his purpose by means of burning candles. Also no one in past ages justified the use of sand in hour glasses by saying that some centuries later interesting laws of motion would be discovered which would give a meaning to the statement that the sand was emptied from the bulbs in equal times. Uniformity in change is directly perceived, and it follows that mankind perceives in nature factors from which a theory of temporal congruence can be formed. The prevalent theory entirely fails to produce such factors.⁷ . . . On the orthodox theory the position of the equations of motion is most ambiguous. The space to which they refer is completely undetermined and so is the measurement of the lapse of time. Science is simply setting out on a fishing expedition to see whether it cannot find some procedure which it can call the measurement of space and some procedure which it can call the measurement of time, and something which it can call a system of forces, and something which it can call masses, so that these formulae may be satisfied. The only reason—on this theory—why anyone should want to satisfy these formulae is a sentimental regard for Galileo, Newton, Euler, and Lagrange. The theory, so far from founding science on a sound observational basis, forces everything to conform to a mere mathematical preference for certain simple formulae.

I do not for a moment believe that this is a true account of the real status of the Laws of Motion. These equations want some slight adjustment for the new formulae of relativity. But

⁷ *Ibid.*, p. 137.

with these adjustments, imperceptible in ordinary use, the laws deal with fundamental physical quantities which we know very well and wish to correlate.

The measurement of time was known to all civilised nations long before the laws were thought of. It is this time as thus measured that the laws are concerned with. Also they deal with the space of our daily life. When we approach to an accuracy of measurement beyond that of observation, adjustment is allowable. But within the limits of observation we know what we mean when we speak of measurements of space and measurements of time and uniformity of change. It is for science to give an intellectual account of what is so evident in sense-awareness. It is to me thoroughly incredible that the ultimate fact beyond which there is no deeper explanation is that mankind has really been swayed by an unconscious desire to satisfy the mathematical formulae which we call the Laws of Motion, formulae completely unknown till the seventeenth century of our epoch.^s

After commenting that purely mathematically, an infinitude of incompatible spatial congruence classes of intervals satisfy the congruence axioms, Whitehead says:

This breakdown of the uniqueness of congruence for space and of its very existence for time is to be contrasted with the fact that mankind does in truth agree on a congruence system for space and a congruence system for time which are founded on the direct evidence of its senses. We ask, why this pathetic trust in the yard measure and the clock? The truth is that we have observed something which the classical theory does not explain.

It is important to understand exactly where the difficulty lies. It is often wrongly conceived as depending on the inexactness of all measurements in regard to very small quantities. According to our methods of observation we may be correct to a hundredth, or a thousandth, or a millionth of an inch. But there is always a margin left over within which we cannot measure. However, this character of inexactness is *not* the difficulty in question.

Let us suppose that our measurements can be ideally exact; it will be still the case that if one man uses one qualifying [i.e., congruence] class γ and the other man uses another qualifying [i.e., congruence] class δ , and if they both admit the

^s *Ibid.*, pp. 139-40.

standard yard kept in the exchequer chambers to be their unit of measurement, they will disagree as to what other distances [at other] places should be judged to be equal to that standard distance in the exchequer chambers. Nor need their disagreement be of a negligible character."⁰ . . .

When we say that two stretches match in respect to length, what do we mean? Furthermore we have got to include time. When two lapses of time match in respect to duration, what do we mean? We have seen that measurement presupposes matching, so it is of no use to hope to explain matching by measurement.¹ . . .

Our physical space therefore must already have a structure and the matching must refer to some qualifying class of qualities inherent in this structure.²

. . . there will be a class of qualities γ one and only one of which attaches to any stretch on a straight line or on a point, such that matching in respect to this quality is what we mean by congruence.

The thesis that I have been maintaining is that measurement presupposes a perception of matching in quality. Accordingly in examining the meaning of any particular kind of measurement we have to ask, What is the quality that matches?³

. . . a yard measure is merely a device for making evident the spatial congruence of the [extended] events in which it is implicated.⁴

Let us begin by inquiring whether Whitehead's historical observation that the human race possessed a time-metric prior to the enunciation of Newton's laws during the seventeenth century can serve to invalidate Poincaré's contentions⁵ that (1) time-congruence in physics is conventional, (2) the definition of temporal congruence used in refined physical theory is given by Newton's laws, and (3) we have no direct intuition of the temporal congruence of non-adjacent time-intervals, the belief in the existence of such an intuition resting on an illusion.

⁰ A. N. Whitehead: *R, op. cit.*, pp. 49–50.

¹ *Ibid.*, pp. 50–51.

² *Ibid.*, p. 51.

³ *Ibid.*, p. 57.

⁴ *Ibid.*, p. 58.

⁵ H. Poincaré: "La Mesure du Temps," *Revue de Métaphysique et de Morale*, Vol. VI (1898), pp. 1–13.

To see how unavailing Whitehead's historical argument is, consider first a hypothetical spatial analogue of his reasoning. We shall see in Chapter Three, Section B that although the demand that Newton's laws be true does uniquely define temporal congruence in the one-dimensional time-continuum, the following *non-analogous* spatial result obtains: the requirement of the Euclideanism of a tabletop does *not* similarly yield a *unique* definition of spatial congruence for that two-dimensional space, there being an infinitude of incompatible congruences issuing in a Euclidean geometry on the tabletop. For the sake of constructing a *hypothetical spatial analogue* to Whitehead's historical argument, however, let us assume that, contrary to fact, it were the case that the requirement of the Euclideanism of the tabletop did uniquely determine the customary definition of rigidity. And now suppose that a philosopher were to say that the latter definition of spatial congruence, like all others, is conventional. What then would be the force of the following kind of putative Whiteheadian assertion: "Well before Hilbert rigorized Euclidean geometry and even much before Euclid less perfectly codified the geometrical relations between the bodies in our environment, men used not only their own limbs but also diverse kinds of solid bodies to certify spatial equality"? Ignoring now refinements required to allow for substance-specific distortions, it is clear that, under the assumed hypothetical conditions, we would be confronted with logically-independent definitions of spatial equality issuing in the same congruence-class of intervals.⁶ The hypothetical concordance of these definitions would indeed be an impressive empirical fact, but, were it to obtain, it could not possibly refute the claim that the one congruence defined alike by all of them is conventional.

Precisely analogous considerations serve to invalidate Whitehead's historical argument regarding time-congruence. Let us grant his implicit assumption of agreement, after allowance for substance-specific idiosyncrasies, between the time congruences defined by various kinds of "clock" devices, so that we can discount hypotheses like that of Milne, which assert the incom-

⁶ In Chapter Four, we shall see in what sense the criterion of rigidity based on the solid body can be regarded as logically-independent of Euclidean geometry when cognizance is taken of substance-specific distortions.

patibility of the congruences defined by "atomic" and "astronomical" clocks. What are some of the clocks adduced by Whitehead as furnishing the "true" time-metric independently of Newton's laws? A candle always made of the same material, of the same size, and having a wick of the same material and size burns very nearly the same number of inches each hour. Hence as early as during the reign of King Alfred (872-900 A.D.), burning candles were used as rough timekeepers by placing notches or marks at such a distance apart that a certain number of spaces would burn each hour.⁷ Ignoring the relatively small variations of the rate of flow of water with the height of the water column in a vessel, the water clock or clepsydra served the ancient Chinese, Byzantines, Greeks, and Romans,⁸ as did the sand-clock, keeping very roughly the same time as burning candles. Again, an essentially frictionless pendulum oscillating with constant amplitude at a point of given latitude on the earth defines the same time-metric as do "natural clocks," i.e., quasi-closed periodic systems.⁹ And, ignoring various refinements, similarly for the rotation of the earth, the oscillations of crystals, the successive round-trips of light over a fixed distance in an inertial system, and the time based on the natural periods of vibrating atoms or "atomic clocks."¹

Thus, unless a hypothesis like that of Milne is right, we find a striking concordance between the time-congruence defined by Newton's amended laws and the temporal equality furnished by several kinds of definitions logically independent of that Newtonian one. This presumed agreement obtains as a matter of presumed *empirical fact* for which the general theory of relativity (hereafter called GTR) has sought to provide an explana-

⁷ W. I. Milham: *Time and Timekeepers* (New York: The Macmillan Co.; 1929), pp. 53-54. A more recent edition appeared in 1941.

⁸ D. J. Price: "The Prehistory of the Clock," *Discovery*, Vol. XVII (1956), pp. 153-57.

⁹ C. Brouwer: "The Accurate Measurement of Time," *Physics Today*, Vol. IV (1951), pp. 7-15.

¹ F. A. B. Ward: *Time Measurement* (4th edition, London: Royal Stationery Office; 1958), Part I, Historical Review; P. Hood: *How Time is Measured* (London: Oxford University Press; 1955); J. J. Baruch: "Horological Accuracy: Its Limits and Implications," *American Scientist*, Vol. XLVI (1958), pp. 188A-196A, and H. Lyons: "Atomic Clocks," *Scientific American*, Vol. CXCVI (1957), pp. 71-82.

of counting was much the same as the one characteristic of chemical reactions: a two-or-threefold increase in rate for a 10°C. rise in temperature. The defenders of the conception that the biological clock is purely internal further adduce observations of the behavior of bees: both outdoors on the surface of the earth and at the bottom of a mine, bees learned to visit at the correct time each day a table on which a dish of syrup was placed daily for a short time at a fixed hour. Since the bees had been found to be hungry for sugar all day long, some investigators hold that neither the assumption that the bees experience periodic hunger, nor the appearance of the sun nor yet the periodicities of the cosmic ray intensity can explain the bees' success in timekeeping. But dosing them with substances like thyroid extract and quinine, which affect the rate of chemical reactions in the body, was found to interfere with their ability to appear at the correct time.⁶

More recently, however, doubt has been cast on the adequacy of the hypothesis of the purely internal clock. A series of experiments with fiddler crabs and other cold-blooded animals⁷ showed that these organisms hold rather precisely to a twenty-four-hour coloration cycle (lightening-darkening rhythm) regardless of whether the temperature at which they are kept is 26 degrees, 16 degrees or 6 degrees Centigrade, although at temperatures near freezing, the color clock changes. It was therefore argued that if the rhythmic timing mechanism were indeed a biochemical one wholly inside the organism, then one would expect the rhythm to speed up with increasing temperature and to slow down with decreasing temperature. And the

⁶ C. S. Pittendrigh and V. G. Bruce: "An Oscillator Model for Biological Clocks," in D. Rudnick (ed.) *Rhythmic and Synthetic Processes in Growth* (Princeton: Princeton University Press; 1957), pp. 75-109, and C. S. Pittendrigh and V. G. Bruce: "Daily Rhythms as Coupled Oscillator Systems and their Relation to Thermoperiodism and Photoperiodism," *Photoperiodism and Related Phenomena in Plants and Animals* (Washington, D.C.: The American Association for the Advancement of Science; 1959).

⁷ F. A. Brown, Jr.: "Biological Clocks and the Fiddler Crab," *Scientific American*, Vol. CXC (1954), pp. 34-37; "The Rhythmic Nature of Animals and Plants," *American Scientist*, Vol. XLVII (1959), pp. 147-68; "Living Clocks," *Science*, Vol. CXXX (1959), pp. 1535-44; and "Response to Pervasive Geophysical Factors and the Biological Clock Problem," *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. XXV (1960), pp. 57-71.

exponents of this interpretation maintain that since the period of the fiddler crab's rhythm remained twenty-four hours through a wide range of temperature, the animals must possess a means of measuring time which is independent of temperature. This, they contend, is "a phenomenon quite inexplicable by any currently known mechanism of physiology, or, in view of the long period-lengths, even of chemical reaction kinetics."⁸ The extraordinary further immunity of certain rhythms of animals and plants to many powerful drugs and poisons which are known to slow down living processes greatly is cited as additional evidence to show that organisms have daily, lunar, and annual rhythms impressed upon them by external physical agencies, thus having access to outside information concerning the corresponding physical periodicities.⁹ The authors of this theory admit, however, that the daily and lunar-tidal rhythms of the animals studied do not depend upon any presently known kind of external cues of the associated astronomical and geophysical cycles.¹ And it is postulated² that these physical cues are being received because living things are able to respond to additional kinds of stimuli at energy levels so low as to have been previously held to be utterly irrelevant to animal behavior. The assumption of such sensitivity of animals is thought to hold out hope for an explanation of animal navigation.

We have dwelled on the two current rival theories regarding the source of the ability of man (and of animals) to make successful estimates of duration introspectively in order to show that, on either theory, the metric of psychological time is tied causally to those physical cycles which serve to define time congruence in physics. Hence when we make the judgment that two intervals of physical time which are equal in the metric of standard clocks also *appear* congruent in the psychometry of mere sense awareness, this justifies only the following innocuous conclusion in regard to physical time: the two intervals in ques-

⁸ F. A. Brown, Jr.: "The Rhythmic Nature of Animals and Plants," *op. cit.*, p. 159.

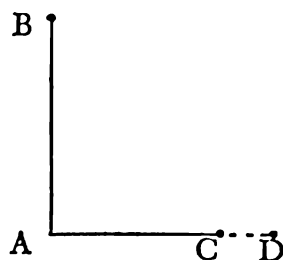
⁹ F. A. Brown, Jr.: "Living Clocks," *op. cit.*, and "Response to Pervasive Geophysical Factors and the Biological Clock Problem," *op. cit.*

¹ F. A. Brown, Jr.: "The Rhythmic Nature of Animals and Plants," *op. cit.*, pp. 153, 166.

² *Ibid.*, p. 168.

tion are congruent by the physical criterion which had furnished the psychometric standard of temporal equality both genetically and epistemologically. How then can the metric deliverances of psychological time possibly show that the time of physics possesses an intrinsic metric, if, as we saw, no such conclusion was demonstrable on the basis of the cycles of physical clocks?

As for *spatial* congruence, we must appraise Whitehead's argument from *matching* in the quotations above after dealing with the following contention by him:³ just as it is an objective datum of experience that two phenomenal color patches have the same color, i.e., are "color-congruent," so also we see that a given rod has the same length in different positions, thus making the latter congruence as objective a relation as the former. Says he: "It is at once evident that all these tests [of congruence by means of steel yard-measures, etc. are] dependent on a direct intuition of permanence."⁴ I take Whitehead to be claiming here that in the accompanying diagram, for example,



the horizontal segment AC could not be stipulated to be congruent to the vertical segment AB. His grounds for this claim are that the deliverances of our visual intuition unequivocally show AC to be shorter than AB, and AB to be congruent to AD, a fact also attested by the finding that a solid rod coinciding with AB to begin with and then rotated into the horizontal position would extend over AC and coincide with AD.

On this my first comment is to ask: what is the significance for the status of the metric of *physical* as distinct from *visual* space of these observational deliverances? And I answer that their significance is entirely consonant with the conventionalist view of physical congruence. The criterion for ocular congruence in

³ A. N. Whitehead: *The Principles of Natural Knowledge*, *op. cit.*, p. 56.

⁴ A. N. Whitehead: *PR*, *op. cit.*, p. 501.

our visual field was presumably furnished both genetically and epistemologically by ocular adaptation to the behavior of transported solids. For when pressed as to what it is about two congruent-looking physical intervals that enables them to sustain the relation of spatial equality, our answer will inevitably have to be: the fact of their capacity to coincide successively with a transported solid rod. Hence when we make the judgment that two intervals of physical space with which transported solid rods coincide in succession also *look* congruent when compared frontally purely by inspection, what this proves in regard to *physical space* is only that these intervals are congruent on the basis of the criterion of congruence which had furnished the genetic and epistemological basis for the ocular congruence to begin with, a criterion given by solid rods. But the visual deliverance of congruence does not constitute an ocular test of the "true" rigidity of solids under transport in the sense of establishing the factuality of the congruence defined by this class of bodies on the basis of an intrinsic metric. Thus, it is a *fact* that in the diagram, *AD* extends over (includes) *AC*, thus being longer. And it will be recalled that Riemann's views on the status of measurement in a spatial continuum require that every definition of "congruent" be consistent with this kind of inclusional fact. How then can visual data possibly interdict our stipulating *AC* to be congruent to *AB* and then allowing for the de facto coincidence of the rotated rod with *AB* and *AD* by assigning to the rod in the horizontal position a length which is suitably *greater* than the one assigned to it in the vertical orientation?

As for his argument from spatial *matching* and its relation to coincidence and measurement, the important issue posed by Whitehead is not whether an operationist account of congruence is adequate. Instead, it is whether spatial congruence derives from intrinsic properties of the intervals concerned rather than wholly from their relation to a transported standard of some kind. For, as we noted earlier, the conventionalist conception of congruence which he is attacking here does not require and is not adequately rendered by the operationist claim that "*the meaning*" of "congruence" is given by some operation of producing coincidence under transport. Just as Einstein's conventionalist view of simultaneity is fully justified only by the ontology of

temporal relatedness postulated by him rather than operationally,⁵ so the conventionalist conception of congruence obtains its philosophical credentials, as Riemann saw, from the assumed continuity of space (and of time). Thus, the question before us again concerns the intrinsicality and uniqueness of congruence, whatever the operations or test conditions by which it may be specified. Hence we again ignore here considerations pertaining to the fact that—on the conventionalist no less than on the Whiteheadian view—congruence is an open-cluster concept in the sense that no one criterion such as coincidence with a transported rod can exhaustively specify its entire actual and potential physical meaning. And then the reply to Whitehead is the following.

If there were a basis for the ascription of an intrinsic metric to space, then he would indeed be entitled to regard coincidence as only a test of congruence *in his sense* that coincidence merely ascertains the equality or matching of the separated intervals concerned with respect to the *intrinsic* amount of space contained by each of them. But without having established the existence of an intrinsic metric, the congruence or *matching* of spatially separated intervals is first *constituted* by the relation which they each bear to the behavior of a transported standard such as a rod or such as the round-trip times of light in an inertial system. And the conventionality of the self-congruence of the latter standard at different places is not at all invalidated by the fact, correctly noted by Whitehead, that measurement presupposes a congruence criterion in terms of which its results are formulated.

That a distance of thirty miles is indeed a long walk for anyone is due to our gait's being tied to the congruence defined by yard (or meter) sticks, thus making it an objective fact that an interval which measures thirty miles in the metric of the yardstick will contain a great many of our steps. But how can this fact possibly show in the absence of an intrinsic metric that the self-congruence of the yardstick under transport is non-conventional?

Moreover, it is plainly false and inconsistent for Whitehead to declare that, according to the classical theory, there is a

⁵ Cf. Chap. Twelve.

"breakdown" of the "very existence" of congruence for time in contrast to the repleteness of space with mutually exclusive congruence classes. For as he himself had noted, on the classical theory a congruence class can be specified for the time continuum by the requirement that Newton's (amended) laws hold. And an infinitude of additional *alternative* time congruences can be given by metrizations based on the values of time variables that are non-linear, one-to-one functions of the Newtonian time variable t .⁶ In his book on Whitehead, Palter is sympathetic to Whitehead to the point of believing that a viable interpretation can generally be given even of those of Whitehead's utterances which are either *prima facie* false or irritatingly obscure. Hence Palter⁷ seeks to defend Whitehead's claim by construing it as referring to the following fact: time, being only a one-dimensional continuum, presents us with no analogue to the higher-dimensional distinction between Euclidean and non-Euclidean geometries, and hence time differs from higher-dimensional space by not possessing a structure corresponding to a distinctive metric geometry. But Palter's defense is unsuccessful, since it rests on the confusion of time's lack of the analogue of a distinctive metric geometry with its alleged lack of a congruence class: the lack of the former does not entail the lack of the latter, although the converse entailment does obtain.

It is significant, however, that there are passages in Whitehead where he comes close to the admission that the pre-eminent role of certain classes of physical objects as our standards of rigidity and isochronism is *not* tantamount to their making evident the *intrinsic* equality of certain spatial and temporal intervals. Thus speaking of the space-time continuum, he says:

This extensive continuum is one relational complex in which all potential objectifications find their niche. It underlies the whole world, past, present, and future. Considered in its full generality, apart from the additional conditions proper only to the cosmic epoch of electrons, protons, molecules, and star-systems, the properties of this continuum are very few and do not include the relationships of metrical geometry.⁸

⁶ For details, see Chap. Two.

⁷ R. M. Palter: *Whitehead's Philosophy of Science*, *op. cit.*, pp. 90–92.

⁸ A. N. Whitehead: *PR*, *op. cit.*, p. 103.

And he goes on to note that there are competing systems of measurement giving rise to alternative families of straight lines and correspondingly alternative systems of metrical geometry of which no one system is more fundamental than any other.⁹ It is in our present cosmic epoch of electrons, protons, molecules, and star-systems that "more special defining characteristics obtain" and that "the ambiguity as to the relative importance of competing definitions of congruence" is resolved in favor of "one congruence definition."² Thus Whitehead maintains that among competing congruence definitions, "That definition which enters importantly into the internal constitutions of the dominating . . . entities is the important definition for the cosmic epoch in question."³ This important concession thus very much narrows the gap between Whitehead's view and the Riemann-Poincaré conception defended in this book: the question as to which of the rival metric geometries is true of physical space is one of objective physical fact only after a congruence definition has been given conventionally by means of the customary rigid body (or otherwise), assuming the usual physical interpretation of the remainder of the geometrical vocabulary.

That the gap between the two views is narrowed by Whitehead's concession here becomes clear upon reading the following statement by him in the light of that concession. Speaking of Sophus Lie's treatment of congruence classes and their associated metric geometries in terms of groups of transformations between points, Whitehead cites Poincaré and says:

The above results, in respect to congruence and metrical geometry, considered in relation to existent space, have led to the doctrine that it is intrinsically unmeaning to ask which system of metrical geometry is true of the physical world. Any one of these systems can be applied, and in an indefinite number of ways. The only question before us is one of convenience in respect to simplicity of statement of the physical laws. This point of view seems to neglect the consideration that science is to be relevant to the definite perceiving minds of men; and that (neglecting the ambiguity introduced by the invariable

⁹ *Ibid.*, p. 149.

¹ *Ibid.*

² *Ibid.*

³ *Ibid.*, p. 506.

slight inexactness of observation which is not relevant to this special doctrine) we have, in fact, presented to our senses a definite set of transformations forming a congruence-group, resulting in a set of measure relations which are in no respect arbitrary. Accordingly our scientific laws are to be stated relevantly to that particular congruence-group. Thus the investigation of the type (elliptic, hyperbolic or parabolic) of this special congruence-group is a perfectly definite problem, to be decided by experiment.⁴

⁴ A. N. Whitehead: *Essays in Science and Philosophy* (New York: The Philosophical Library; 1947), p. 265.

Chapter 2

THE SIGNIFICANCE OF ALTERNATIVE TIME METRIZATIONS IN NEWTONIAN MECHANICS AND IN THE GENERAL THEORY OF RELATIVITY

(A) NEWTONIAN MECHANICS

On the conception of time congruence as conventional, the preference for the customary definition of isochronism—a preference not felt by Einstein in the general theory of relativity (GTR), as we shall see in Section B—can derive only from considerations of convenience and elegance so long as the resulting form of the theory is not prescribed. Hence, the thesis that isochronism is conventional precludes a difference in factual import (content) or in *explanatory power* between two descriptions one of which employs the customary isochronism while the other is a “translation” (transcription) of it into a language employing a time congruence incompatible with the customary one. As a test case for this thesis of *explanatory parity*, the general outline of a counter-argument has been suggested which we shall be able to state after some preliminaries.

On the Riemannian analysis, congruence must be regarded as conventional in the time continuum of Newtonian dynamics no less than in the theory of relativity. We shall therefore wish to compare in regard to explanatory capability the two forms of Newtonian dynamics corresponding to two different time congruences as follows.

The first of these congruences is defined by the requirement that Newton's laws hold, as modified by the addition of very small corrective terms expressing the so-called relativistic motion of the perihelia. This time congruence will be called "Newtonian," and the time variable whose values represent Newtonian time after a particular unit has been chosen will be denoted by "t." The second time congruence is defined by the rotational motion of the earth. It does not matter for our purpose whether we couple the latter congruence with a unit given by the mean solar second, which is the 1/86400 part of the mean interval between two consecutive meridian passages of the fictitious mean sun, or with a different unit given by the sidereal day, which is the interval between successive meridian passages of a star. What matters is that both the mean solar second and the sidereal day are based on the periodicities of the earth's rotational motion. Assume now that one or another of these units has been chosen, and let T be the time variable associated with that metrization, which we shall call "diurnal time." The important point is that the time variables t and T are non-linearly related and are associated with *incompatible* definitions of isochronism, because the speed of rotation of the earth *varies relatively to the Newtonian time-metric* in several distinct ways.¹ Of these, the best known is the relative slowing down of the earth's rotation by the tidal friction between the water in the shallow seas of the earth and the land under it. Upon calculating the positions of the moon, for example, via the usual theory of celestial mechanics, which is based on the Newtonian time-metric, the observed positions of the moon in the sky would be found to be *ahead* of the calculated ones *if* we were to identify the time defined by the earth's rotation with the Newtonian time of celestial mechanics. And the same is true of the positions of the planets of the solar system and of the moons of Jupiter in amounts all corresponding to a slowing down on the part of the earth.

Now consider the following argument for the lack of explanatory parity between the two forms of the dynamical theory respectively associated with the t- and T-scales of time: "*Dynamical*

¹ G. M. Clemence: "Time and its Measurement," *American Scientist*, Vol. XL (1952), pp. 264-67.

facts will discriminate in favor of the t-scale as opposed to the T-scale. It is granted that it is *kinematically* equivalent to say

- (a) the earth's rotational motion has slowed down relatively to the "clocks" constituted by various revolving planets and satellites of the solar system,

or

- (b) the revolving celestial bodies speed up their periodic motions relatively to the earth's uniform rotation.

But these two statements are not on a par explanatorily in the context of the dynamical theory of the motions in the solar system. For whereas the slowing down of the earth's rotation in formulation (a) *can* be understood as the dynamical effect of nearby masses (the tidal waters and their friction), no similar dynamical *cause* can be supplied for the accelerations in formulation (b). And the latter fact shows that a theory incorporating formulation (a) has greater explanatory power or factual import than a theory containing (b).² In precisely this vein, d'Abro, though stressing on the one hand that apart from convenience and simplicity there is nothing to choose between different metrics,² on the other hand adduces the provision of *causal* understanding by the t-scale as an argument in its favor and thus seems to construe such differences of simplicity as involving factually *non-equivalent* descriptions:

if in mechanics and astronomy we had selected at random some arbitrary definition of time, if we had defined as congruent the intervals separating the rising and setting of the sun at all seasons of the year, say for the latitude of New York, our understanding of mechanical phenomena would have been beset with grave difficulties. As measured by these new temporal standards, free bodies would no longer move with constant speeds, but would be subjected to periodic accelerations *for which it would appear impossible to ascribe any definite cause*, and so on. As a result, the law of inertia would have to be abandoned, and with it the entire doctrine of classical mechanics, together with Newton's law. Thus a change in our understanding of congruence would entail far-reaching consequences.

² A. d'Abro: *The Evolution of Scientific Thought from Newton to Einstein*, *op. cit.*, p. 53.

Again, in the case of the vibrating atom, had some arbitrary definition of time been accepted, we should have had to assume that the same atom presented the most capricious frequencies. *Once more it would have been difficult to ascribe satisfactory causes* to these seemingly haphazard fluctuations in frequency; and a simple understanding of the most fundamental optical phenomena would have been well-nigh impossible.³

To examine this argument, let us set the two formulations of dynamics corresponding to the t - and T -scales respectively before us mathematically in order to have a clearer statement of the issue.

The differences between the two kinds of temporal congruence with which we are concerned arise from the fact that the functional relationship

$$T = f(t)$$

relating the two time scales is *non-linear*, so that time-intervals which are congruent on the one scale are generally incongruent on the other. It is clear that this function is monotone-increasing, and thus we know that permanently

$$\frac{dT}{dt} \neq 0.$$

Moreover, in view of the non-linearity of $T = f(t)$, we know that dT/dt is *not* constant. Since the function f has an inverse, it will be possible to translate any set of laws formulated on the basis of either of the two time-scales into the corresponding other scale. In order to see what form the customary Newtonian force law assumes in diurnal time, we must express the acceleration ingredient in that law in terms of diurnal time. But in order to derive the transformation law for the accelerations, we first treat the velocities. By the chain rule for differentiation, we have, using " r " to denote the position vector,

$$(1) \quad \frac{dr}{dt} = \frac{dr}{dT} \frac{dT}{dt}.$$

Suppose a body is *at rest* in the coordinate system in which r is measured, *when Newtonian time is employed*; then this body

³ *Ibid.*, p. 78, my italics.

will also be held to be *at rest diurnally*: since we saw that the second term on the right-hand side of equation (1) cannot be zero, the left-hand side of (1) will vanish if and only if the first term on the right-hand side of (1) is zero. Though rest in a given frame in the *t*-scale will correspond to rest in that frame in the *T*-scale as well, equation (1) shows that the *constancy* of the non-vanishing Newtonian velocity dr/dt will *not* correspond to a constant diurnal velocity dr/dT , since the derivative dT/dt changes with both Newtonian and diurnal time. Now, differentiation of equation (1) with respect to the Newtonian time *t* yields

$$(2) \quad \frac{d^2r}{dt^2} = \frac{dr}{dT} \frac{d^2T}{dt^2} + \frac{dT}{dt} \frac{d}{dt} \left(\frac{dr}{dT} \right).$$

But, applying the chain-rule to the second factor in the second-term on the right-hand side of (2), we obtain

$$(2a) \quad \frac{d}{dt} \left(\frac{dr}{dT} \right) = \frac{d^2r}{dT^2} \frac{dT}{dt}.$$

Hence (2) becomes

$$(3) \quad \frac{d^2r}{dt^2} = \frac{dr}{dT} \frac{d^2T}{dt^2} + \frac{d^2r}{dT^2} \left(\frac{dT}{dt} \right)^2.$$

Solving for the diurnal acceleration, and using equation (1) as well as the abbreviations

$$f'(t) \equiv \frac{dT}{dt} \text{ and } f''(t) \equiv \frac{d^2T}{dt^2}, \text{ we find}$$

$$(4) \quad \underbrace{\frac{d^2r}{dT^2}}_{\text{diurnal acceleration}} = \underbrace{\frac{1}{[f'(t)]^2} \frac{d^2r}{dt^2}}_{\text{Newtonian acceleration}} - \overbrace{\frac{f''(t)}{[f'(t)]^3} \frac{dr}{dt}}^{\text{secular term}} \underbrace{\frac{dr}{dt}}_{\text{Newtonian velocity}}.$$

Several ancillary points should be noted briefly in regard to equation (4) before seeing what light it throws on the form assumed by causal explanation within the framework of a diurnal description. When the Newtonian force on a body is *not* zero because the body is accelerating under the influence of masses,

the diurnal acceleration will generally also not be zero, save in the unusual case when

$$(5) \quad \frac{d^2r}{dt^2} = \frac{f''(t)}{f'(t)} \frac{dr}{dt}.$$

Thus the causal influence of masses, which gives rise to the Newtonian accelerations in the usual description, is seen in (4) to make a definite contribution to the diurnal acceleration as well. But the new feature of the diurnal description of the facts lies in the possession of a *secular acceleration* by all bodies not at rest, even when no masses are inducing Newtonian accelerations, so that the first term on the right-hand side of (4) vanishes. And this secular acceleration is numerically *not* the same for all bodies but depends on their velocities dr/dt in the given reference frame and thus also on the reference frame.

The character and existence of this secular acceleration calls for several kinds of comment.

Its dependence on the velocity and on the reference frame should neither occasion surprise nor be regarded as a difficulty of any sort. As to the velocity-dependence of the secular acceleration, consider a simple numerical example which removes any surprise: if instead of calling two successive hours on Big Ben equal, we remetrized time so as to assign the measure one half hour to the second of these intervals, then all bodies having uniform speeds in the usual time metric will double their speeds on the new scale after the first interval, and the *numerical* increase or *acceleration* in the speeds of initially faster bodies will be greater than that in the speeds of the initially slower bodies. Now as for the dependence of the secular acceleration on the reference frame, in the context of the physical facts asserted by the Newtonian theory *apart from* its metrical philosophy, it is a mere prejudice to require that, to be admissible, every formulation of that theory must agree with the customary one in making the acceleration of a body at any given time be the same in all Galilean reference frames ("Galilean relativity"). For not a single bona fide physical fact of the Newtonian world is overlooked or contradicted by a kinematics not featuring this Galilean relativity. It is instructive to be aware in this connection that even in the *customary* rendition of the *kinematics of special*

relativity, a constant acceleration in a frame S' would not generally correspond to a constant acceleration in a frame S , because the component accelerations in S depend not only on the accelerations in S' but also on the component velocities in that system which would be changing with the time.

But what are we to say, apart from the dependence on the velocity and reference system, about the very presence of this "dynamically-unexplained" or causally baffling *secular acceleration*? To deal with this question, we first observe merely for comparison that in the *customary* formulation of Newtonian mechanics, constant *speeds* (as distinct from constant *velocities*) fall into *two* classes with respect to being attributable to the dynamical action of perturbing masses: constant *rectilinear* speeds are affirmed to prevail in the absence of any mass influences, while constant *curvilinear* (e.g., circular) speeds are related to the (centripetally) accelerating actions of masses. Now in regard to the presence of a secular acceleration in the diurnal description, it is fundamental to see the following: Whereas on the version of Newtonian mechanics employing the *customary* metrizations (of time and space), *all* accelerations whatsoever in Galilean frames are of *dynamical* origin by being attributable to the action of specific masses, *this feature of Newton's theory is made possible not only by the facts but also by the particular time-metrization chosen to codify them*. As equation (4) shows upon equating d^2r/dt^2 to zero, the dynamical character of *all* accelerations is not vouchsafed by any causal facts of the world with which every theory would have to come to terms. For the diurnal description encompasses the objective behavior of bodies (point-events and coincidences) as a function of the presence or absence of other bodies no less than does the Newtonian one, thereby achieving full explanatory parity with the latter in all logical (as distinct from pragmatic!) respects.

Hence the provision of a dynamical basis for all accelerations should not be regarded as an *inflexible epistemological requirement* in the elaboration of a theory explaining mechanical phenomena. Disregarding the pragmatically decisive consideration of convenience, there can therefore be no valid explanatory objection to the diurnal description, in which accelerations fall into *two* classes by being the superpositions, in the sense of equation

(4), of a dynamically grounded *and* a kinematically grounded quantity. And, most important, since there is no *slowing down* of the earth's rotation on the diurnal metric, there can be no question in that description of specifying a *cause* for such a non-existent deceleration; instead, a frictional *cause* is now specified for the earth's diurnally-*uniform* rotation *and* for the liberation of heat accompanying this kind of uniform motion. For in the T-scale description, it is *uniform* rotation which requires a dynamical cause constituted by masses interacting (frictionally) with the uniformly rotating body, and it is now a law of nature or entailed by such a law that all diurnally-uniform rotations issue in the dissipation of heat. Of course, the mathematical representation of the frictional interaction will not have the customary Newtonian form: to obtain the diurnal account of the frictional dynamics of the tides, one would need to apply transformations of the kind given in our equation (4) to the quantities appearing in the relevant Newtonian equations for this case.⁴

But, it will be asked, what of the Newtonian conservation principles, if the T-scale of time is adopted? It is readily demonstrable by reference to the simple case of the motion of a free particle that while the Newtonian kinetic energy will be constant in this case, its formal diurnal homologue (as opposed to its diurnal equivalent!) will *not* be constant. Let us denote the constant *Newtonian* velocity of the free particle by " v_t ," the subscript "t" serving to represent the use of the t-scale, and let " v_T " denote the diurnal velocity corresponding to v_t . Since we know from equation (1) above that

$$v_t = v_T \frac{dT}{dt},$$

where v_t is constant but dT/dt is *not*, we see that the diurnal *homologue* $\frac{1}{2}mv_T^2$ of the Newtonian kinetic energy cannot be constant in this case, although the diurnal *equivalent* $\frac{1}{2}m(v_T \frac{dT}{dt})^2$ of the constant Newtonian kinetic energy $\frac{1}{2}mv_t^2$ is necessarily constant. Just as in the case of the Newtonian equations of motion themselves, so also in the case of the Newtonian conservation

⁴ For these equations, cf. H. Jeffreys: *The Earth* (3rd ed.; Cambridge: Cambridge University Press; 1952), Chap. 8, and G. I. Taylor: "Tidal Friction in the Irish Sea," *Philosophical Transactions of the Royal Society, A.*, Vol. CCXX (1920), pp. 1-33.

principle of mechanical energy, the diurnal equivalent or transcription explains all the facts explained by the Newtonian original. Hence our critic can derive no support at all from the fact that the formal diurnal homologues of Newtonian conservation principles generally do not hold. And we see, incidentally, that the time-invariance of a physical quantity and hence the appropriateness of singling it out from among others as a form of "energy," etc., will depend not only on the facts but also on the time-metrization used to render them. It obviously will not do, therefore, to charge the diurnal description with inconsistency via the *petitio* of grafting onto it the requirement that it incorporate the homologues of Newtonian conservation principles which are incompatible with it: a case in point is the charge that the diurnal description violates the conservation of energy because in its metric the frictional generation of heat in the tidal case is not compensated by any reduction in the speed of the earth's rotation! Whether the diurnal time-metrization permits the deduction of conservation principles of a *relatively simple type* involving diurnally based quantities is a rather involved mathematical question whose solution is not required to establish our thesis that, apart from pragmatic considerations, the diurnal description enjoys explanatory parity with the Newtonian one.

We have been disregarding pragmatic considerations in assessing the explanatory capabilities of two descriptions associated with different time-metrizations as to parity. But it would be an error to infer that in pointing to the equivalence of such descriptions in regard to factual content, we are committed to the view that there is no criterion for choosing between them and hence no reason for preferring any one of them to the others. Factual adequacy (truth) is, of course, the cardinal *necessary* condition for the acceptance of a scientific theory, but it is hardly a *sufficient* condition for accepting any one particular formulation of it which satisfies this necessary condition. As well say that a person pointing out that equivalent descriptions can be given in the decimal (metric) and English system of units cannot give telling reasons for preferring the former! Indeed, after first commenting on the *factual basis* for the existence of the Newtonian time congruence, we shall see that there are weighty *pragmatic* reasons for preferring that metrization of the time continuum. And these reasons will turn out to be entirely consonant with our

twin contention that alternative metrization allows linguistically different, equivalent descriptions *and* that geo-chronometric conventionalism is *not* a subthesis of TSC.

The *factual basis* for the existence of the Newtonian time-metrization will be appreciated by reference to the following two considerations: first, as we shall prove presently, it is a highly fortunate empirical fact, and not an a priori truth, that there exists a time-metrization *at all* in which *all* accelerations with respect to inertial systems are of dynamic origin, as claimed by the Newtonian theory, and second, it is a further empirical fact that the time-metrization having this remarkable property (i.e., "ephemeris time") is furnished physically by the earth's annual revolution around the sun (not by its diurnal rotation) albeit not in any observationally simple way, since due account must be taken computationally of the irregularities produced by the gravitational influences of the other planets.⁵ That the existence of a time-metrization in which *all* accelerations with respect to inertial systems are of *dynamical* origin cannot be guaranteed a priori is demonstrable as follows.

Suppose that, *contrary to actual fact*, it *were* the case that a *free* body did accelerate with respect to an inertial system when its motion is described in the metric of ephemeris time t , it thus being assumed that there are accelerations in the customary time metric which are *not* dynamical in origin. More particularly, let us now posit that, contrary to actual fact, a *free* particle were to execute *one-dimensional* simple harmonic motion of the form

$$r = \cos \omega t,$$

where r is the distance from the origin. In that hypothetical eventuality, the acceleration of a *free* particle in the t -scale would have the time-dependent value

$$\frac{d^2r}{dt^2} = -\omega^2 \cos \omega t.$$

And our problem is to determine whether there would then exist some other time-metrization $T = f(t)$ possessing the Newtonian

⁵ G. M. Clemence: "Astronomical Time," *Reviews of Modern Physics*, Vol. XXIX (1957), pp. 2-8; "Dynamics of the Solar System," ed. by E. Condon and H. Odishaw, *Handbook of Physics* (New York: McGraw-Hill Book Company, Inc.; 1958), p. 65; "Ephemeris Time," *Astronomical Journal*, Vol. LXIV (1959), pp. 113-15 and *Transactions of the International Astronomical Union*, Vol. X (1958); "Time and Its Measurement," *op. cit.*

property that our *free* particle has a *zero* acceleration. We shall now find that the answer is definitely negative: under the hypothetical empirical conditions which we have posited, there would indeed be no admissible single-valued time-metrization T at all in which all accelerations with respect to an inertial system would be of dynamical origin.

For let us now regard T in equation (4) of this chapter as the time variable associated with the sought-after metrization $T = f(t)$ in which the acceleration d^2r/dT^2 of our free particle would be *zero*. We recall that equation (5) of this chapter was obtained from equation (4) by equating the T -scale acceleration d^2r/dT^2 to zero. Hence *if* our sought-after metrization exists at all, it would have to be the solution $T = f(t)$ of the scalar form of equation (5) as applied to our one-dimensional motion. That equation is

$$(6) \quad \frac{d^2r}{dt^2} = \frac{f''(t)}{f'(t)} \frac{dr}{dt}.$$

Putting $v \equiv dr/dt$ and noting that

$$\frac{d}{dt} \log f'(t) = \frac{f''(t)}{f'(t)}, \text{ and } \frac{d}{dt} \log v = \frac{1}{v} \frac{dv}{dt},$$

equation (6) becomes

$$\frac{d}{dt} \log v = \frac{d}{dt} \log f'(t).$$

Integrating, and using $\log c$ as the constant of integration, we obtain

$$\log v = \log cf'(t),$$

or

$$v = cf'(t),$$

which is

$$\frac{dr}{dt} = c \frac{dT}{dt}.$$

Integration yields

$$(7) \quad r = cT + d,$$

where d is a constant of integration. But, by our earlier hypothesis, $r = \cos \omega t$. Hence (7) becomes

$$(8) \quad T = \frac{1}{c} \cos \omega t - \frac{d}{c}.$$

It is evident that the solution $T = f(t)$ given by equation (8) is *not* a *one-to-one* function: the *same* time T in the sought-after metrization would correspond to all those *different* times on the t -scale at which the oscillating particle would return to the *same place* $r = \cos \omega t$ in the course of its periodic motion. And by thus violating the basic topological requirement that the function $T = f(t)$ be one-to-one, the T -scale which does have the sought-after Newtonian property under our *hypothetical* empirical conditions is physically a quite inadmissible and hence unavailable metrization.

It follows that there is no a priori assurance of the existence of at least one time-metrization possessing the Newtonian property that the acceleration of a free particle with respect to inertial systems is zero. So much for the *factual basis* of the existence of the Newtonian time-metrization.

Now, inasmuch as the employment of the time-metrization based on the earth's annual revolution issues in Newton's relatively simple laws, there are powerful reasons of mathematical tractability and convenience for greatly preferring the time-metrization in which *all* accelerations with respect to inertial systems are of dynamical origin. In fact, the various refinements which astronomers have introduced in their physical standards for temporal congruence have been dictated by the demand for a definition of temporal congruence (or of a so-called "invariable" time standard) for which Newton's laws will hold in the solar system, including the relatively simple conservation laws interconnecting diverse kinds of phenomena (mechanical, thermal, etc.). And thus, as Feigl and Maxwell have aptly put it, one of the important criteria of descriptive simplicity which greatly restrict the range of "reasonable" conventions is seen to be the *scope* which a convention will allow for mathematically tractable laws.

(B) THE GENERAL THEORY OF RELATIVITY

See Append. §5

In the special theory of relativity, only the *customary* time-metrization is employed in the following sense: At any given point A in a Galilean frame, the length of a time-interval between two events at the point A is given by the *difference*

between the time-coordinates of the two events as furnished by the readings of a standard clock at A whose periods are defined to be congruent. This is, of course, the precise analogue of the customary definition of spatial congruence which calls the rod congruent to itself everywhere, when at relative rest, after allowance for substance-specific perturbations. On the other hand, as we shall now see, there are contexts in which the *general* theory of relativity (GTR) utilizes a criterion of *temporal* congruence which is an analogue of a *non*-customary kind of spatial congruence in the following sense: the length of a time-interval separating two events at a clock depends not only on the *difference* between the time-coordinates which the clock assigns to these events but also on the *spatial* location of the clock (though not on the time itself at which the time-interval begins or ends).

A case in point from the GTR involves a *rotating disk* to which we apply those principles which the GTR takes over from the special theory of relativity. Let a set of standard material clocks be distributed at various points on such a disk. The infinitesimal application of the special relativity clock retardation then tells us the following: a clock at the *center* O of the disk will maintain the rate of a contiguous clock located in an inertial system I with respect to which the disk has angular velocity ω , *but* the same does not hold for clocks located at other points A of the disk which are at positive distances r from O. Such A-clocks have various linear velocities ωr relatively to I in virtue of their common angular velocity ω . Accordingly, all A-clocks (whatever their chemical constitution) will have readings lagging behind the corresponding readings of the respective I-system clocks adjacent to them by a *factor* of $\sqrt{1 - \frac{r^2\omega^2}{c^2}}$, where c is the velocity of light. What would be the consequence of using the *customary* time-metrization everywhere on the rotating disk and letting the duration (length) of a time-interval elapsing at a given point A be given by the *difference* between the time-coordinates of the termini of that interval as furnished by the readings of the standard clock at A? The adoption of the customary time-metric would saddle us with a *most complicated* description of the propagation of light in the rotating system having the following undesirable features: (i) time would

enter the description of nature explicitly in the sense that the one-way velocity of light would depend on the time, since the lagging *rate* of the clock at A issues in a temporal change in the magnitude of the one-way transit-time of a light ray for journeys between O and A, and (ii) the number of light waves emitted at A during a unit of time on the A-clock is *greater* than the number of waves arriving at the center O in one unit of time on the O-clock.⁶ To avoid the undesirably complicated laws entailed by the use of the simple customary definition of time congruence, the GTR jettisoned the latter. In its stead, it adopted the following more complicated, non-customary congruence definition for the sake of the simplicity of the resulting laws: at any point A on the disk the length (duration) of a time-interval is given *not* by the *difference* between the A-clock coordinates of its termini but by the product of this increment *and* the rate factor $\frac{1}{\sqrt{1 - \frac{r^2\omega^2}{c^2}}}$, which depends

on the spatial coordinate r of the point A. This rate factor serves to assign a *greater* duration to time-intervals than would be obtained from the customary procedure of letting the length of time be given by the increment in the clock readings. In view of the *dependence* of the metric on the *spatial position* r , via the rate factor entering into it, we are confronted here with a *non-customary* time-metrization fully as consonant with the temporal *order* of the events at A as is the customary metric.

A similarly non-standard time-metric is used by Einstein in his tentative GTR paper of 1911⁷ in treating the effect of gravitation on the propagation of light. Analysis shows that the very same complexities in the description of light propagation which are encountered on the rotating disk arise here as well, if the standard time-metric is used. These complexities are eliminated here in quite analogous fashion by the use of a non-customary time-metric. Thus, if we are concerned with light emitted on

⁶ C. Møller: *The Theory of Relativity* (Oxford: Oxford University Press; 1952), pp. 225–26.

⁷ A. Einstein: "On the Influence of Gravitation on the Propagation of Light," in *The Principle of Relativity*, a collection of memoirs (New York: Dover Publications, Inc.; 1953), Sec. 3.

the sun and reaching the earth and if " $-\Phi$ " represents the negative difference in gravitational potential between the sun and the earth, then we proceed as follows: prior to being brought from the earth to the sun, a clock is set to have a rate *faster* than that of an adjoining terrestrial clock by a factor of $\frac{1}{1 - \frac{\Phi}{c^2}}$ (to a first approximation), where $\frac{\Phi}{c^2} < 1$.

Chapter 3

CRITIQUE OF REICHENBACH'S AND CARNAP'S PHILOSOPHY OF GEOMETRY

(A) THE STATUS OF "UNIVERSAL FORCES"

In *Der Raum*,¹ Carnap begins his discussion of *physical* space by inquiring whether and how a line in this space can be identified as straight. Arguing from *testability* and not, as we did in Chapter One, from the continuity of that manifold, he answers this inquiry as follows: "It is impossible in principle to ascertain this, if one restricts oneself to the unambiguous deliverances of experience and does not introduce freely chosen conventions in regard to objects of experience."² And he then points out that the most important convention relevant to whether certain physical lines are to be regarded as straights is the specification of the metric ("Mass-setzung"), which is conventional because it could "never be either confirmed or refuted by experience." Its statement takes the following form: "A particular body and two fixed points on it are chosen, and it is then agreed what length is to be assigned to the interval between these points under various conditions (of temperature, position, orientation, pressure, electrical charge, etc.). An example of the choice of a metric is the stipulation that the two marks on the Paris standard meter bar define an interval of $100 \cdot f(T; \phi, \lambda, h; \dots)$ cm; . . . a

¹ R. Carnap: *Der Raum* (Berlin: Reuther and Reichard, 1922), p. 33.

² *Ibid.*

unit must also be chosen, but that is not our concern here which is with the choice of the body itself and with the function $f(T, \dots)$.³

Once a particular function f has been chosen, the coincidence behavior of the selected transported body permits the determination of the metric tensor g_{ik} appropriate to that choice, thereby yielding a *congruence* class of intervals and the associated geometry. Accordingly, Carnap's thesis is that the question as to the geometry of physical space is indeed an *empirical* one but subject to an important proviso: it becomes empirical only *after* a physical definition of congruence for line segments has been given *conventionally* by stipulating (to within a constant factor depending on the choice of unit) what length is to be assigned to a transported solid rod in different positions of space.

Like Carnap, Reichenbach invokes testability⁴ to defend this qualified empiricist conception of geometry and speaks of "the relativity of geometry"⁵ to emphasize the dependence of the geometry on the definition of congruence. Carnap had lucidly conveyed the conventionality of congruence by reference to our freedom to choose the function f in the metric. But Reichenbach couches this conception in *metaphorical* terms by speaking of "universal forces"⁶ whose metrical "effects" on measuring rods are then said to be a matter of convention as follows: the customary definition of congruence in which the rod is held to be of equal length everywhere (*after* allowance for substance-specific thermal effects and the like) corresponds to equating the universal forces to zero; on the other hand, a non-customary definition of congruence, according to which the length of the rod varies with position or orientation (even *after* allowance for thermal effects, etc.), corresponds to assuming an appropriately specified non-vanishing universal force whose mathematical characterization will be explained below. Reichenbach did not anticipate that this metaphorical encumbrance of his formulation would mislead some people into making the ill-conceived charge that non-customary definitions of congruence

³ *Ibid.*, pp. 33–34.

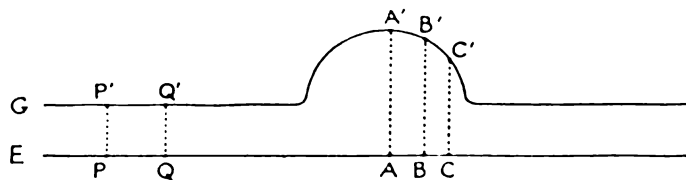
⁴ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 16.

⁵ *Ibid.*, Sec. 8.

⁶ *Ibid.*, Secs. 3, 6, 8.

are based on *ad hoc* invocations of universal forces. Inasmuch as this charge has been leveled against the conventionality of congruence, it is essential that we now divest Reichenbach's statement of its misleading potentialities.

Reichenbach⁷ invites consideration of a large hemisphere made of glass which merges into a huge glass plane, as shown in cross-section by the surface G in the diagram, which consists of a plane with a hump. Using solid rods, human beings on this surface would readily determine it to be a Euclidean plane with a central hemispherical hump. He then supposes an opaque



plane E to be located below the surface G as shown in the accompanying figure. Vertical light rays incident upon G will cast shadows of all objects on that glass surface onto E. As measured by *actual* solid rods, G-people will find A'B' and B'C' to be equal, while their projections AB and BC on the Euclidean plane E would be unequal. Reichenbach now wishes to prepare the reader for the recognition of the conventionality of congruence by having him deal with the following kind of question. Might it not be the case that:

(1) the *inequality* of AB and BC is only *apparent*, these intervals and other projections like them in the region R of E under the hemisphere being *really* equal, so that the *true* geometry of the surface E is *spherical* in R and Euclidean only outside it,

(2) the equality of A'B' and B'C' is only *apparent*, the *true* geometry of surface G being plane Euclidean *throughout*, since in the apparently hemispherical region R' of G *real* equality obtains among those intervals which are the upward vertical projections of E-intervals in R that are equal in the *customary* sense of our daily life, and

(3) on each of the two surfaces, transported measuring rods respectively fail to coincide with really equal intervals in R and

⁷ *Ibid.*, Sec. 3.

R' respectively, because they do not remain truly congruent to themselves under transport, being deformed under the influence of undetectable forces which are "universal" in the sense that (a) they affect all materials alike, and (b) they penetrate all insulating walls?

On the basis of the conceptions presented in Chapter One, *which involve no kind of reference to universal forces*, one can fulfill Reichenbach's desire to utilize this question as a basis for expounding the conventionality of congruence by presenting the following considerations. The legitimacy of making a distinction between the real (true) and the apparent geometry of a surface *turns on the existence of an intrinsic metric*. If there were an intrinsic metric, there would be a basis for making the distinction between real (true) and apparent equality of a rod under transport, and thereby between the true and the apparent geometry. But inasmuch as there is not, the question as to whether a given surface is really a Euclidean plane with a hemispherical hump or only apparently so must be *replaced* by the following question: on a *particular convention of congruence* as specified by a choice of one of Carnap's functions f , does the coincidence behavior of the transported rod on the surface in question yield the geometry under discussion or not?

Thus the question as to the geometry of a surface is inherently ambiguous without the introduction of a congruence definition. And in view of the conventionality of spatial congruence, we are entitled to metrize G and E *either* in the customary way *or* in other ways so as to describe E as a Euclidean plane with a hemispherical hump R in the center and G as a Euclidean plane throughout. To assure the correctness of the latter *non*-customary descriptions, we need only *decree* the congruence of those respective intervals which our questioner called "really equal" as opposed to apparently equal in parts (1) and (2) in his question respectively. Accordingly, without the presupposition of an intrinsic metric there can be no question of an absolute or "real" deformation of *all* kinds of measuring rods alike under the action of universal forces, and, *mutatis mutandis*, the same statement applies to clocks. Since a rod undergoes no kind of objective physical change in the supposed "presence" of universal forces, that "presence" signifies no more than that we assign a different

length to it in different positions or orientations by convention. Hence, just as the conversion of the length of a table from meters to feet does not involve the action of a force on the table as the "cause" of the change, so also reference to universal forces as "causes" of "changes" in the transported rod can have no literal but only *metaphorical* significance. Moreover, mention of universal forces is an entirely dispensable mode of speech in this context as is evident from the fact that the rule assigning to the transported rod lengths which vary with its position and orientation can be given by specifying Carnap's function f .

Reichenbach, on the other hand, chooses to formulate the conventionality of congruence by first distinguishing between what he calls "differential" and "universal" forces and then using "universal forces" *metaphorically* in his statement of the epistemological status of the metric. By "differential" forces⁸ he means thermal and other influences which we called "perturbational" in Chapter One and whose presence is distorting by issuing in the dependence of the coincidence behavior of transported rods on the latter's chemical composition. Since we conceive of physical geometry as the system of metric relations which are *independent* of chemical composition, we correct for the substance-specific deformations induced by differential forces.⁹ Reichenbach defines "universal forces" as having the twin properties of affecting all materials in the same way and as all-permeating because there are no walls capable of providing insulation against them. There is precedent for a literal rather than metaphorical use of universal forces to give a congruence definition: in order to provide a *physical realization* of a non-customary congruence definition which would metrize the interior of a sphere of radius R so as to be a model of an infinite 3-dimensional hyperbolic space, Poincaré¹ postulates that (a) each concentric sphere of radius $r < R$ is held at a constant

⁸ *Ibid.*

⁹ *Ibid.*, p. 26; H. P. Robertson: "Geometry as a Branch of Physics," *Albert Einstein: Philosopher-Scientist*, ed. by P. A. Schilpp (Evanston: Library of Living Philosophers; 1949), pp. 327-29, and "The Geometries of the Thermal and Gravitational Fields," *American Mathematical Monthly*, Vol. LVII (1950), pp. 232-45; and E. W. Barankin: "Heat Flow and Non-Euclidean Geometry," *American Mathematical Monthly*, Vol. XLIX (1942), pp. 4-14.

¹ H. Poincaré: *The Foundations of Science*, *op. cit.*, pp. 75-77.

absolute temperature $T \propto (R^2 - r^2)$, while the optical index of refraction is inversely proportional to $R^2 - r^2$, and (b) *contrary to actual fact*, all kinds of bodies within the sphere have the *same* coefficient of thermal expansion. It is essential to see that the expansions and contractions of these bodies under displacement have a *literal* meaning in this context, because they are *relative* to the actual displacement-behavior of our normally Euclidean kinds of bodies and are linked to thermal sources.²

But Reichenbach's *metaphorical* use of universal forces for giving the congruence definition and exhibiting the dependence of the geometry on that definition takes the following form: "Given a geometry G' to which the measuring instruments conform [*after* allowance for the effects of thermal and other 'differential' influences], we can imagine a universal force F which affects the instruments in such a way that the actual geometry is an arbitrary geometry G , while the observed deviation from G is due to a universal deformation of the measuring instruments."³ And he goes on to say that if g'_{ik} ($i = 1, 2, 3; k = 1, 2, 3$) are the empirically obtained metrical coefficients of the geometry G' and g_{ik} those of G , then the force tensor F is given mathematically by the tensor equation

$$g'_{ik} + F_{ik} = g_{ik},$$

where the g'_{ik} , which had yielded the observed geometry G' , are furnished experimentally by the measuring rods⁴ and where the F_{ik} are the "correction factors" $g_{ik} - g'_{ik}$ which are added correctionally to the g'_{ik} so that the g_{ik} are obtained.⁵ But since

² A precisely analogous *literal* use of universal forces is also made by Reichenbach [H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, pp. 11-12] to convey pictorially a physical realization of a congruence definition which would confer a spherical geometry on the region R of the surface E discussed above.

³ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 33.

⁴ For details on this experimental procedure, see, for example, *ibid.*, Secs. 39 and 40.

⁵ We shall see in Section (ii) of this chapter that Reichenbach was *mistaken* in asserting [*ibid.*, pp. 33-34] that for a given surface or 3-space a particular metric geometry *determines* (1) a *unique* definition of congruence and, once a unit of length has been chosen, (2) a *unique* set of functions g_{ik} as the representations of the metric tensor in any particular coordinate system. It will turn out that there are infinitely many *incompatible* congruence definitions and as many correspondingly different metric tensors

Reichenbach emphasizes that it is a matter of *convention* whether we equate F_{ik} to zero or not.⁶ This formulation is merely a *metaphorical* way of asserting that the following is a matter of convention: whether congruence is said to obtain among intervals having equal lengths ds given by the metric $ds^2 = g'_{ik}dx^i dx^k$ —which entails G' as the geometric description of the observed coincidence relations—or among intervals having equal lengths ds given by the metric $ds^2 = g_{ik}dx^i dx^k$ which yields a different geometry G .⁷ Clearly then, to equate the universal forces to zero is merely to choose the metric based on the tensor g'_{ik} which was obtained from measurements in which the rod was called congruent to itself everywhere. In other words, to stipulate $F_{ik} = 0$ is to choose the customary congruence standard based on the rigid body.

which impart the *same* geometry to physical space. Hence, while a given metric tensor yields a unique geometry, a geometry G does not determine a metric tensor *uniquely* to within a constant factor depending on the choice of unit. And thus it is incorrect for Reichenbach to speak here of the components g_{ik} of a particular metric tensor as "those of G " [*ibid.*, p. 33n] and to suppose that a unique F is specified by the requirement that a certain geometry G prevail in place of the observed geometry G' .

⁶ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, pp. 16, 27–28, 33.

⁷ It is to be clearly understood that the g_{ik} yield a congruence relation incompatible with the one furnished by the g'_{ik} , because in any given coordinate system they are different functions of the given coordinates and not proportional to one another. (A difference consisting in a mere proportionality could not involve a difference in the congruence classes but only in the unit of length used.) The incompatibility of the congruences furnished by the two sets of metric coefficients is a necessary though not a sufficient condition (cf. the preceding footnote 5) for the non-identity of the associated geometries G and G' .

The difference between the two metric tensors corresponding to *incompatible* congruences must not be confounded with a mere difference in the *representations* in different coordinate systems of the *one* metric tensor corresponding to a *single* congruence criterion (for a given choice of a unit of length): the former is illustrated by the *incompatible* metrizations $ds^2 = dx^2 + dy^2$ and $ds^2 = (dx^2 + dy^2)/y^2$ in which the corresponding metric coefficients are suitably *different* functions of the *same* rectangular coordinates, whereas the latter is illustrated by using first rectangular and then polar coordinates to express the *same metric* as follows: $ds^2 = dx^2 + dy^2$ and $ds^2 = d\rho^2 + \rho^2 d\theta^2$. In the latter case, we are not dealing with different *metrizations* of the space but only with different *coordinatizations* (*parametrizations*) of it, at least one pair of corresponding metric coefficients having members which are *different* functions of their respective coordinates but so chosen as to yield an *invariant* ds .

On the other hand, apart from one exception to be stated presently, to stipulate that the components F_{ik} may not all be zero is to adopt a non-customary metric given by a tensor g_{ik} corresponding to a specified variation of the length of the rod with position or orientation.

That there is one exception, which, incidentally, Reichenbach does not discuss, can be seen as follows: a given congruence determines the metric tensor up to a constant factor depending on the choice of unit and conversely. Hence two metric tensors will correspond to different congruences if and only if they differ other than by being proportional to one another. Thus, if g_{ik} and g'_{ik} are proportional by a factor different from 1, these two tensors furnish metrics differing only in the choice of unit and hence yield the same congruence. Yet in the case of such proportionality of g_{ik} and g'_{ik} , the F_{ik} *cannot* all be zero. For example, if we consider the line element

$$ds^2 = a^2 d\phi^2 + a^2 \sin^2 \phi d\theta^2$$

on the surface of a sphere of radius $a = 1 \text{ meter} = 100 \text{ cm}$, the mere change of units from meters to centimeters would change the metric $ds^2 = d\phi^2 + \sin^2 \phi d\theta^2$ into

$$ds^2 = 10000 d\phi^2 + 10000 \sin^2 \phi d\theta^2.$$

And if these metrics are identified with g'_{ik} and g_{ik} respectively, we obtain

$$F_{11} = g_{11} - g'_{11} = 10000 - 1 = 9999$$

$$F_{12} = F_{21} = g_{12} - g'_{12} = 0$$

$$F_{22} = g_{22} - g'_{22} = 10000 \sin^2 \phi - \sin^2 \phi = 9999 \sin^2 \phi.$$

It is now apparent that the F_{ik} , being given by the differences between the g_{ik} and the g'_{ik} , will not all vanish *and* that also these two metric tensors will yield the *same* congruence, if and only if these tensors are proportional by a factor different from 1. Therefore, a necessary and sufficient condition for obtaining *incompatible* congruences is that there be at least *one* non-vanishing component F_{ik} such that the metric tensors g_{ik} and g'_{ik} are not proportional to one another. The exception to our statement that a non-customary congruence definition is assured by the failure of at least one component of the F_{ik} to vanish is

therefore given by the case of the proportionality of the metric tensors.⁸

Although Reichenbach's metaphorical use of "universal forces" issued in misleading and wholly unnecessary complexities which we shall point out presently, he himself was in no way victimized by them. Writing in 1951 concerning the invocation of universal forces in this context, he declared: "The assumption of such forces means merely a change in the coordinative definition of congruence."⁹ It is therefore quite puzzling that, in 1956, Carnap, who had lucidly expounded the same ideas *non-metaphorically* as we saw, singled out Reichenbach's characterization and recommendation of the customary congruence definition in terms of equating the universal forces to zero as a praiseworthy part of Reichenbach's outstanding work *The Philosophy of Space and Time*. In his Preface to the latter work, Carnap says:

Of the many fruitful ideas which Reichenbach contributed . . . I will mention only one, which seems to me of great interest for the methodology of physics but which has so far not found the attention it deserves. This is the principle of the elimination of universal forces. . . . Reichenbach proposes to accept as a general methodological principle that we choose that form of a theory among physically equivalent forms (or, in other words, that definition of "rigid body" or "measuring standard") with respect to which all universal forces disappear.¹

The misleading potentialities of reference to metaphorical "universal forces" in the statement of the congruence definition and in related contexts manifest themselves in three ways as follows:

(1) The formulation of non-customary congruence definitions in terms of deformations by universal forces has inspired the erroneous charge that such congruences are *ad hoc* because they

⁸ There is a most simple illustration of the fact that if the metric tensors are *not* proportional such that as few as *one* of the components F_{ik} is *non-vanishing*, the congruence associated with the g_{ik} will be incompatible with that of the g'_{ik} and will hence be non-customary. If we consider the metrics $ds^2 = dx^2 + dy^2$ and $ds^2 = 2dx^2 + dy^2$, then comparison of an interval for which $dx = 1$ and $dy = 0$ with one for which $dx = 0$ and $dy = 1$ will yield congruence on the first of these metrics but not on the second.

⁹ H. Reichenbach: *The Rise of Scientific Philosophy* (Berkeley: University of California Press; 1951), p. 133.

¹ R. Carnap: Preface to H. Reichenbach's *The Philosophy of Space and Time* (New York: Dover Publications, Inc.; 1957), p. 7.

allegedly involve the *ad hoc* postulation of non-vanishing universal forces.

(2) In Reichenbach's statement of the congruence definition to be employed to explore the spatial geometry in a *gravitational* field, universal forces enter in both a *literal* and a *metaphorical* sense. The conflation of these two senses issues in a seemingly contradictory formulation of the customary congruence definition.

(3) Since the variability of the curvature of a space would manifest itself in alterations in the coincidence behavior of all kinds of solid bodies under displacement, Reichenbach speaks of bodies displaced in such a space as being subject to universal forces "destroying coincidences."² In a manner analogous to the gravitational case, the conflation of this literal sense with the metaphorical one renders the definition of rigidity for this context paradoxical.

We shall now briefly discuss these three sources of confusion in turn.

1. If a congruence definition itself had factual content, so that alternative congruences would differ in factual content, then it would be significant to say of a congruence definition that it is *ad hoc* in the sense of being an *evidentially-unwarranted* claim concerning facts. But inasmuch as ascriptions of spatial congruence to non-coinciding intervals are not factual but conventional, neither the customary nor any of the non-customary definitions of congruence can possibly be *ad hoc*. Hence the abandonment of the former in favor of the latter kind of definition can be no more *ad hoc* than the regraduation of a Centigrade thermometer into a Fahrenheit one or than the change from Cartesian to polar coordinates. By formulating non-customary congruence definitions in terms of the metaphor of universal forces, Reichenbach made it possible for his metaphorical sense to be misconstrued as literal. And once this error had been committed, its victims tacitly regarded the customary congruence definition as factually true and felt justified in dismissing other congruences as *ad hoc* on the grounds that they involved the *ad hoc* assumption of (literally conceived) universal forces. As well say that a change of the units of length is *ad hoc*.

² H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 27.

Thus we find that Ernest Nagel, for example, overlooks in writing about Poincaré that the invocation of universal forces to preserve Euclidean geometry can be no more *ad hoc* than a change from rectangular to polar coordinates in order to write the equation of a circle as $\rho = k$ instead of $\sqrt{x^2 + y^2} = k$. After granting that, if necessary, Euclidean geometry can be retained by an appeal to universal forces, Nagel writes: "Nevertheless, universal forces have the curious feature that their presence can be recognized *only* on the basis of geometrical considerations. The assumption of such forces thus has the appearance of an *ad hoc* hypothesis, adopted solely for the sake of salvaging Euclid."³ But it is only Nagel's and not Poincaré's construal of the character of the relevant kind of universal forces which permits Nagel's conclusion that Poincaré must have recourse to some sort of *ad hoc* hypothesis in order to assure a Euclidean description of the posited observational facts. For within the framework of a physical theory which assumes space to be a mathematical continuum—an assumption on which Poincaré's entire thesis was predicated—there is no basis whatever for leveling an *ad hoc* charge against Poincaré. The invocation of the kind of universal force whose "presence can be recognized *only* on the basis of geometrical considerations" and which is introduced "solely for the sake of salvaging Euclid" can be no more *ad hoc* than the use of polar rather than rectangular coordinates. Nagel's concern that Poincaré's thesis can be upheld only on pain of an *ad hoc* assumption of universal forces is just as unwarranted as the following supposition: the numerical increase in the lengths of all objects produced by the conversion from meters to inches requires the *ad hoc* postulation of universal forces as the "physical cause" of the universal elongation.

2. Regarding the geometry in a gravitational field, Reichenbach says the following: "We have learned . . . about the difference between *universal* and *differential* forces. These concepts have a bearing upon this problem because we find that gravitation is a universal force. It does indeed affect all bodies in the same manner. This is the physical significance of the equality of gravitational and inertial mass."⁴ It is entirely correct, of course,

³ E. Nagel: *The Structure of Science* (New York: Harcourt, Brace and World; 1961), p. 264.

⁴ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 256.

that a uniform gravitational field (which has not been transformed away in a given space-time coordinate system) is a universal force in the *literal* sense *with respect to a large class of effects* such as the free fall of bodies. But there are other effects, such as the bending of elastic beams, with respect to which gravity is clearly a *differential* force in Reichenbach's sense: a wooden book shelf will sag more under gravity than a steel one. And this shows, incidentally, that *Reichenbach's classification of forces into universal and differential is not mutually exclusive*. Of course, just as in the case of any other force having differential effects on measuring rods, allowance is made for differential effects of gravitational origin in laying down the congruence definition.

The issue is therefore twofold: first, does the fact that gravitation is a universal force in the *literal* sense indicated above have a bearing on the spatial geometry, and second, in the presence of a gravitational field is the logic of the spatial congruence definition any different in regard to the role of *metaphorical* universal forces from what it is in the absence of a gravitational field? Within the particular context of the GTR, there is indeed a *literal* sense in which the gravitational field of the sun, for example, is causally relevant geometrically as a universal force. And the literal sense in which the coincidence behavior of the transported customary rigid body is objectively different in the vicinity of the sun, for example, from what it is in the absence of a gravitational field can be expressed in two ways as follows: (1) *relatively to the congruence defined by the customary rigid body*, the spatial geometry in the gravitational field is *not* Euclidean—contrary to pre-GTR physics—but is Euclidean in the absence of a gravitational field; (2) the geometry in the gravitational field is Euclidean if and only if the customary congruence definition is supplanted by one in which the length of the rod varies *suitably* with its position or orientation,⁵ whereas it is Euclidean relatively to the customary congruence definition for a vanishing gravitational field. It will be noted, however, that formulation (1) makes no mention at all of any

⁵ For the gravitational field of the sun, the function specifying a non-customary congruence definition issuing in a Euclidean geometry is given in R. Carnap: *Der Raum*, *op. cit.*, p. 58.

deformation of the rod by universal forces as that body is transported from place to place in a given gravitational field. Nor need there be any metaphorical reference to universal forces in the statement of the customary congruence definition ingredient in formulation (1). For that statement can be given as follows: in the presence no less than in the absence of a gravitational field, congruence is conventional, and hence we are free to adopt the customary congruence in a gravitational field as a basis for determining the spatial geometry. By encumbering his statement of a congruence definition with metaphorical use of "universal forces," Reichenbach enables the unwary to infer incorrectly that a rod subject to the universal force of gravitation in the specified *literal* sense *cannot* consistently be regarded as *free* from deforming universal forces in the *metaphorical* sense and hence cannot serve as the congruence standard. This conflation of the literal and metaphorical senses of "universal force" in the context of a theory which assumes the continuity of space thus issues in the mistaken belief that in the GTR the customary spatial congruence definition cannot be adopted consistently for the gravitational field. And those who were led to this misconception by Reichenbach's metaphor will therefore deem self-contradictory the following consistent assertion by him: "We do not speak of a change produced by the gravitational field in the measuring instruments, but regard the measuring instruments as 'free from deforming forces' in spite of the gravitational effects."⁶ Moreover, those victimized by the metaphorical part of Reichenbach's language will be driven to reject as inconsistent Einstein's characterization of the geometry in the gravitational field in the GTR as given in formulation (1) above. And they will insist erroneously that formulations (1) and (2) are *not* equally acceptable alternatives on the grounds that formulation (2) is uniquely correct.

The confounding of the literal and metaphorical senses of "universal force" by reference to gravity in a theoretical context which presupposes the continuity of space is present, for example, in Ernest Nagel's treatment of universal forces with resulting potentialities of confusion. Thus, he incorrectly cites the force

⁶ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 256.

that a uniform gravitational field (which has not been transformed away in a given space-time coordinate system) is a universal force in the *literal* sense *with respect to a large class of effects* such as the free fall of bodies. But there are other effects, such as the bending of elastic beams, with respect to which gravity is clearly a *differential* force in Reichenbach's sense: a wooden book shelf will sag more under gravity than a steel one. And this shows, incidentally, that *Reichenbach's classification of forces into universal and differential is not mutually exclusive*. Of course, just as in the case of any other force having differential effects on measuring rods, allowance is made for differential effects of gravitational origin in laying down the congruence definition.

The issue is therefore twofold: first, does the fact that gravitation is a universal force in the *literal* sense indicated above have a bearing on the spatial geometry, and second, in the presence of a gravitational field is the logic of the spatial congruence definition any different in regard to the role of *metaphorical* universal forces from what it is in the absence of a gravitational field? Within the particular context of the GTR, there is indeed a *literal* sense in which the gravitational field of the sun, for example, is causally relevant geometrically as a universal force. And the literal sense in which the coincidence behavior of the transported customary rigid body is objectively different in the vicinity of the sun, for example, from what it is in the absence of a gravitational field can be expressed in two ways as follows: (1) *relatively to the congruence defined by the customary rigid body*, the spatial geometry in the gravitational field is *not* Euclidean—contrary to pre-GTR physics—but is Euclidean in the absence of a gravitational field; (2) the geometry in the gravitational field is Euclidean if and only if the customary congruence definition is supplanted by one in which the length of the rod varies *suitably* with its position or orientation,⁵ whereas it is Euclidean relatively to the customary congruence definition for a vanishing gravitational field. It will be noted, however, that formulation (1) makes no mention at all of any

⁵ For the gravitational field of the sun, the function specifying a non-customary congruence definition issuing in a Euclidean geometry is given in R. Carnap: *Der Raum*, *op. cit.*, p. 58.

deformation of the rod by universal forces as that body is transported from place to place in a given gravitational field. Nor need there be any metaphorical reference to universal forces in the statement of the customary congruence definition ingredient in formulation (1). For that statement can be given as follows: in the presence no less than in the absence of a gravitational field, congruence is conventional, and hence we are free to adopt the customary congruence in a gravitational field as a basis for determining the spatial geometry. By encumbering his statement of a congruence definition with metaphorical use of "universal forces," Reichenbach enables the unwary to infer incorrectly that a rod subject to the universal force of gravitation in the specified *literal* sense *cannot* consistently be regarded as *free* from deforming universal forces in the *metaphorical* sense and hence cannot serve as the congruence standard. This conflation of the literal and metaphorical senses of "universal force" in the context of a theory which assumes the continuity of space thus issues in the mistaken belief that in the GTR the customary spatial congruence definition cannot be adopted consistently for the gravitational field. And those who were led to this misconception by Reichenbach's metaphor will therefore deem self-contradictory the following consistent assertion by him: "We do not speak of a change produced by the gravitational field in the measuring instruments, but regard the measuring instruments as 'free from deforming forces' in spite of the gravitational effects."⁶ Moreover, those victimized by the metaphorical part of Reichenbach's language will be driven to reject as inconsistent Einstein's characterization of the geometry in the gravitational field in the GTR as given in formulation (1) above. And they will insist erroneously that formulations (1) and (2) are *not* equally acceptable alternatives on the grounds that formulation (2) is uniquely correct.

The confounding of the literal and metaphorical senses of "universal force" by reference to gravity in a theoretical context which presupposes the continuity of space is present, for example, in Ernest Nagel's treatment of universal forces with resulting potentialities of confusion. Thus, he incorrectly cites the force

⁶ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 256.

of gravitation in its role of being *literally* a "universal force" as a *species* of what are only *metaphorically* "universal forces" within a theory that assumes the continuity of physical space. Specifically, in speaking of Poincaré's assumption of universal forces whose "presence can be recognized *only* on the basis of geometrical considerations" because they are assumed "solely for the sake of salvaging Euclid"—i.e., "universal forces" in the *metaphorical* sense—Nagel says: "'Universal force' is not to be counted as a 'meaningless' phrase, for it is evident that a procedure is indicated for ascertaining whether such forces are present or not. Indeed, gravitation in the Newtonian theory of mechanics is just such a universal force; it acts alike on all bodies and cannot be screened."⁷ But Nagel's misidentification of Newtonian gravity as "just such a universal force" as the metaphorical kind of universal force whose "presence can be recognized *only* on the basis of geometrical considerations" can lead to the incorrect conclusion that a rod subject to a Newtonian gravitational field *cannot* be held to be "free" from "universal forces" in the sense of being self-congruent when performing its metrical function under transport.

Precisely the latter kind of error was committed by P. K. Feyerabend, who discusses the bearing of a hypothetical universal force of the *literal* kind on the metrical function of a transported rod. For Feyerabend supposes mistakenly that the rod must be held to be *distorted* in the presence of universal forces of the *literal* kind which "act upon all chemical substances in a similar way but which make themselves noticed in a slight change induced in the transitions probabilities of atoms radiating in that region."⁸

3. In a manner analogous to the gravitational case just discussed, we can assert the following: since congruence is conventional, we are at liberty to use the customary definition of it without regard to whether the geometry obtained by measurements expressed in terms of that definition is one of *variable*

⁷ E. Nagel: *The Structure of Science* (New York: Harcourt, Brace and World; 1961), p. 264, n. 19.

⁸ Cf. P. K. Feyerabend: "Comments on Grünbaum's 'Law and Convention in Physical Theory,'" in H. Feigl and G. Maxwell (eds.) *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart and Winston; 1961), p. 157, and A. Grünbaum: "Rejoinder to Feyerabend," *ibid.*, pp. 164–67.

curvature or not. And thus we see that upon avoiding the intrusion of a metaphorical use of "universal forces," the statement of the congruence definition need not take any cognizance of whether the resulting geometry will be one of constant curvature or not. A geometry of constant curvature or so-called "congruence geometry" is characterized by the fact that the so-called "axiom of free mobility" holds in it: for example, on the surface of a sphere, a triangle having certain angles and sides in a given place can be moved about without any change in their magnitudes relatively to the customary standards of congruence for angles and intervals. By contrast, on the surface of an egg, the failure of the axiom of free mobility to hold can be easily seen from the following indicator of the variability of the curvature of that 2-space: a circle and its diameter made of any kind of wire are so constructed that one end P of the diameter is attached to the circle while the other end S is free *not* to coincide with the opposite point Q on the circle though coinciding with it in a given initial position on the surface of the egg. Since the ratio of the diameter and the circumference of a circle *varies* in a space of variable curvature such as that of the egg-surface, S will no longer coincide with Q if the circular wire and its attachment PS are moved about on the egg so as to preserve contact with the egg-surface everywhere. The indicator thus exhibits an objective destruction of the coincidence of S and Q which is wholly independent of the indicator's chemical composition (under uniform conditions of temperature, etc.).

One may therefore speak *literally* here, as Reichenbach does,⁹ of universal forces acting on the indicator by *destroying coincidences*. And since the customary congruence definition is entirely permissible as a basis for geometries of variable curvature, there is, of course, no inconsistency in giving that congruence definition by equating universal forces to *zero* in the *metaphorical* sense, even though the destruction of coincidences attests to the quite *literal* presence of causally efficacious universal forces. But Reichenbach invokes universal forces *literally* without any warning of an impending metaphorical reference to them in the congruence definition. And the reader is therefore both startled and

⁹ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, Sec. 6.

puzzled by the seeming paradox in Reichenbach's declaration that: "Forces *destroying coincidences* must also be set equal to zero, if they satisfy the properties of the universal forces mentioned on p. 13; only then is the problem of geometry uniquely determined."¹ Again the risk of confusion can be eliminated by dispensing with the metaphor in the congruence definition.

Although I believe that Reichenbach's *The Philosophy of Space and Time* is the most penetrating single book on its subject, the preceding analysis shows why I cannot share Nagel's judgment that, in that book, Reichenbach employs "The distinction between 'universal' and 'differential' forces . . . with great clarifying effect."²

Having divested Reichenbach's statements about universal forces of their misleading potentialities, we shall hereafter safely be able to discuss other issues raised by statements of his which are couched in terms of universal forces.

The first of these is posed by the following assertion by him: "We obtain a statement about physical reality only if in addition to the geometry *G* of the space its universal field of force *F* is specified. Only the combination

$$G + F$$

is a testable statement."³ In order to be able to appraise this assertion, consider a surface on which some set of generalized curvilinear (or "Gaussian") coordinates has been introduced. The coordinatization of a space, whose purpose it is merely to number points so as to convey their topological neighborhood relations of betweenness, does not, *as such*, presuppose (entail) a metric. However, the statement of a *rule* assuring that different people will independently effect the *same* coordinatization of a given space may require reference to the use of a rod. But even in the case of coordinates such as rectangular (Cartesian) coordinates, whose assignment is carried out by the use of a rigid rod, it is quite possible to ignore the manner in which the coordinatization was effected and to regard the coordinates purely topologically, so that a metric very *different* from $ds^2 = dx^2 + dy^2$ can then still be introduced quite consistently. Accordingly, now put a metric

¹ *Ibid.*, p. 27.

² E. Nagel: *The Structure of Science*, *op. cit.*, p. 264, n. 18.

³ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 33.

$ds^2 = g_{ik}dx^i dx^k$ onto the coordinated surface quite arbitrarily by a capricious choice of a suitable set of functions g_{ik} of the given coordinates. Assume that the latter specification of the geometry G is not coupled with any information regarding F .

Is it then correct to say that since this metrization provides no information at all about the coincidence behavior of a rod under transport on the surface, it conveys no factual information whatever about the surface or physical reality? That such an inference is mistaken can be seen from the following: depending upon whether the Gaussian curvature K associated with the stipulated g_{ik} is positive (spherical geometry), zero (Euclidean geometry), or negative (hyperbolic geometry), it is an objective fact about the surface that through a point outside a given geodesic of the chosen metric, there will be respectively 0, 1, or infinitely many other such geodesics which will not *intersect* the given geodesic. Whether or not certain lines on a surface intersect, is however merely a *topological* fact concerning it. And hence we can say that although an arbitrary metrization of a space without a specification of F is not altogether devoid of factual content pertaining to that space, such a metrization can yield no objective facts concerning the space not already included in the latter's topology.

We can therefore conclude the following: if the description of a space (surface) is to contain empirical information concerning the coincidence behavior of transported rods in that space and if a metric $ds^2 = g_{ik}dx^i dx^k$ (and thereby a geometry G) is chosen whose congruences do not accord with those defined by the application of the transported rod, then indeed Reichenbach's assertion holds. Specifically, the chosen metric tensor g_{ik} and its associated geometry G must then be coupled with a specification of the different metric tensor g'_{ik} that would have been found experimentally, if the rod had actually been chosen as the congruence standard. But Reichenbach's provision of that specification via the universal force F is quite unnecessarily roundabout. For F is defined by $F_{ik} = g_{ik} - g'_{ik}$ and cannot be known without already knowing *both* metric tensors. Thus, there is a loss of clarity in pretending that the metric tensor g'_{ik} , which codifies the empirical information concerning the rod, is first obtained from the identity

$$g'_{ik} = g_{ik} - F_{ik}.$$

(B) THE "RELATIVITY OF GEOMETRY"

In order to emphasize the dependence of the metric geometry on the definition of congruence, Reichenbach speaks of "the relativity of geometry." But there is an important error in his characterization of this dependence, which takes the following forms: "If we change the coordinative definition of congruence, a different geometry will result. This fact is called the *relativity of geometry*,"⁴ and, more explicitly, "There is nothing wrong with a coordinative definition established on the requirement that a certain kind of geometry is to result from the measurements. . . . A coordinative definition can also be introduced by the prescription what the result of the measurements is to be. 'The comparison of length is to be performed in such a way that Euclidean geometry will be the result'—this stipulation is a possible form of a coordinative definition."⁵ Reichenbach's claim is that a given metric geometry *uniquely* determines a congruence class (or congruence definition) appropriate to it.

That this contention is mistaken will now be demonstrated: we shall show that besides the customary definition of congruence, which assigns the same length to the measuring rod everywhere and thereby confers a Euclidean geometry on an ordinary tabletop, there are infinitely many *other* definitions of congruence having the following property: they likewise yield a Euclidean geometry for that surface but are incompatible with the customary one by making the length of a rod depend on its orientation and/or position.

Thus, consider our horizontal tabletop equipped with a network of Cartesian coordinates x and y , but now metrize this surface by means of the *non-standard* metric

$$ds^2 = \sec^2\theta dx^2 + dy^2,$$

where $\sec^2\theta$ is a *constant greater than 1*. Unlike the standard metric, this metric assigns to an interval whose coordinates differ by dx *not* the length dx but the *greater* length $\sec\theta dx$ while continuing to assign the length dy to an interval whose coordinates differ only by dy . Although this metric thereby makes the

⁴ H. Reichenbach: *The Rise of Scientific Philosophy*, *op. cit.*, p. 132.

⁵ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, pp. 33–34

length of a given rod dependent on its orientation, we shall show that the *infinitely* many different non-standard congruences generated by values of $\sec\theta$ which exceed 1 each impart a *Euclidean* geometry to the tabletop no less than does the standard congruence given by

$$ds^2 = dx^2 + dy^2.$$

Accordingly, our demonstration will show that the requirement of Euclideanism does *not* uniquely determine a congruence class of intervals but allows an *infinitude* of *incompatible* congruences. We shall therefore have established that there are infinitely many ways in which a measuring rod could squirm under transport on the tabletop as compared to its familiar de facto behavior while still yielding a Euclidean geometry for that surface.

To carry out the required demonstration, we first note the preliminary fact that the geometry yielded by a particular metrization is clearly independent of the particular coordinates in which that metrization is expressed. And hence if we expressed the standard metric

$$ds^2 = dx^2 + dy^2$$

in terms of the primed coordinates x' and y' given by the transformations

$$\begin{aligned} x &= x' \sec\theta \\ y &= y', \end{aligned}$$

obtaining

$$ds^2 = \sec^2\theta \, dx'^2 + dy'^2$$

we would obtain a *Euclidean* geometry as before, since the latter equation would merely express the original standard metric in terms of the primed coordinates. Thus, when the same *invariant* ds of the standard metric is expressed in terms of both primed and unprimed coordinates, the metric coefficients g'_{ik} given by $\sec^2\theta$, 0 and 1 yield a Euclidean geometry no less than do the unprimed coefficients 1, 0, and 1.

This elementary ancillary conclusion now enables us to see that the following *non*-standard metrization (or remetrization) of the surface in terms of the *original*, unprimed rectangular coordinates must likewise give rise to a Euclidean geometry:

$$ds^2 = \sec^2\theta \, dx^2 + dy^2.$$

For the value of the Gaussian curvature and hence the prevailing geometry depends not on the particular coordinates (primed or unprimed) to which the metric coefficients g_{ik} pertain but only on the *functional form* of the g_{ik} ⁶ which is the same here as in the case of the g'_{ik} above.

More generally, therefore, the geometry resulting from the *standard* metrization is *also* furnished by the following kind of *non-standard* metrization of a space of points, expressed in terms of the same (unprimed) coordinates as the standard one: the *non-standard* metrization has unprimed metric coefficients g_{ik} which have the same *functional form* (to within an arbitrary constant arising from the choice of unit of length) as those primed coefficients g'_{ik} which are obtained by expressing the *standard* metric in some set or other of *primed* coordinates via a suitable coordinate transformation. In view of the large variety of allowable coordinate transformations, it follows at once that the class of *non-standard* metrizations yielding a Euclidean geometry for a tabletop is far wider than the already infinite class given by

$$ds^2 = \sec^2\theta \, dx^2 + dy^2, \text{ where } \sec^2\theta > 1.$$

Thus, for example, there is *identity of functional form* between the *standard* metric in *polar* coordinates, which is given by

$$ds^2 = d\rho^2 + \rho^2 d\theta^2,$$

and the *non-standard* metric in Cartesian coordinates given by

$$ds^2 = dx^2 + x^2 dy^2,$$

since x plays the same role *formally* as ρ , and similarly for y and θ . Consequently, the latter *non-standard* metric issues in a *Euclidean* geometry just as the former standard one does.

It is clear that the multiplicity of metrizations which we have proven for Euclidean geometry obtains as well for each of the non-Euclidean geometries. The failure of a geometry of two or more dimensions to determine a congruence definition uniquely does not, however, have a counterpart in the *one-dimensional* time-continuum: the demand that Newton's laws hold in their

⁶ F. Klein: *Vorlesungen über Nicht-Euklidische Geometrie* (Berlin: Springer-Verlag, 1928), p. 281.

customary metrical form does determine a unique definition of temporal congruence. And hence it is feasible to rely on the law of translational or rotational inertia to define a time metric or "uniform time."

On the basis of this result, we can now show that a number of claims made by Reichenbach and Carnap respectively are false.

(1) In 1951, Reichenbach wrote, as will be recalled from the beginning of this Section: "If we change the coordinative definition of congruence, a different geometry will result. This fact is called the *relativity of geometry*."⁷ That this statement is false is evident from the fact that if, in our example of the tabletop, we change our congruence definition from $ds^2 = dx^2 + dy^2$ to any one of the infinitely many definitions incompatible with it that are given by $ds^2 = \sec^2\theta \, dx^2 + dy^2$, precisely the same Euclidean geometry results. Thus, contrary to Reichenbach, the introduction of a non-vanishing universal force corresponding to an alternative congruence does not guarantee a change in the geometry. Instead, the correct formulation of the relativity of geometry is that in the form of the ds function the congruence definition uniquely determines the geometry, though not conversely, and that any one of the congruence definitions issuing in a geometry G' can always be replaced by infinitely many *suitably* different congruences yielding a specified *different* geometry G . In view of the unique fixation of the geometry by the congruence definition in the context of the facts of coincidence, the repudiation of a given geometry in favor of a different one does indeed require a change in the definition of congruence. And the new congruence definition which is expected to furnish the new required geometry can do so in one of the following two ways:

(1) by determining a system of geodesics *different* from the one yielded by the original congruence definition, or

(2) if the geodesics determined by the new congruence definition are the *same* as those associated with the original definition, then the *angle* congruences must be different,⁸ i.e., the new

⁷ H. Reichenbach: *The Rise of Scientific Philosophy*, op. cit., p. 132.

⁸ The specification of the magnitudes assigned to angles by the components g_{ik} of the metric tensor was given in Chapter One, Section C.

congruence definition will have to require a different congruence class of angles.⁹

That (2) constitutes a genuine possibility for obtaining a different geometry will be evident from the following example of mapping the sphere geodesically on a plane, a mapping which issues in two *incompatible* definitions of congruence

$$\begin{aligned} ds^2_1 &= g_{ik}dx^i dx^k, \text{ and} \\ ds^2_2 &= g'_{ik}dx^i dx^k \end{aligned}$$

that yield the *same* system of geodesics via the equations $\delta \int ds_1 = 0$ and $\delta \int ds_2 = 0$ and yet determine *different* geometries (Gaussian curvatures), because they require incompatible congruence classes of *angles* appropriate to these respective geometries. A horizontal surface which is a Euclidean plane on the *customary* metrization can alternatively be metrized to have the geometry of a hemisphere by projection from the *center* of a sphere through its *lower half* while the south pole is resting on that plane. Upon calling congruent on the horizontal surface segments and angles which are the projections of equal segments and angles respectively on the lower hemisphere, the great circle arcs of the hemisphere map into the Euclidean straight lines of the plane such that *every* straight of the Euclidean description is *also* a straight (geodesic) of the new hemispherical geometry conferred on the horizontal surface.¹ But the *angles* which are regarded as congruent on the horizontal surface in the new metrization are not congruent in the original metrization yielding a Euclidean description. The destruction by the new metrization of the angle congruences associated with the original metrization can be made apparent as follows.

Consider two triangles ABC and A'B'C' which qualify as *similar* in the Euclidean geometry of the original metrization, so that $\sphericalangle A = \sphericalangle A'$, $\sphericalangle B = \sphericalangle B'$, and $\sphericalangle C = \sphericalangle C'$. Since the geodesics of

⁹ For other general theorems governing the so-called "geodesic correspondence" or "geodesic mapping" relevant here, cf. L. P. Eisenhart: *An Introduction to Differential Geometry* (Princeton: Princeton University Press; 1947), Sec. 37, pp. 205–11, and D. J. Struik: *Differential Geometry*, *op. cit.*, pp. 177–80.

¹ The mathematical details can be found in D. J. Struik: *Differential Geometry*, *op. cit.*, p. 179.

the new metrization, i.e., of its associated non-Euclidean (spherical) geometry are the *same* as those of Euclidean geometry in this case, triangles ABC and A'B'C' will still be *rectilinear* triangles on the new metrization. But since there are no similar triangles in the spherical geometry of constant positive Gaussian curvature which results from the new metrization, triangles ABC and A'B'C' are no longer similar on the new metrization though still rectilinear. Hence the original angle congruences among these triangles can no longer obtain on the new metrization.

It must be pointed out, however, that if a change in the congruence definition *preserves* the geodesics, then its issuance in a *different* congruence class of *angles* is only a necessary and not a sufficient condition for imparting to the surface a metric geometry *different* from the one yielded by the original congruence definition. This fact becomes evident by reference to our earlier case of the tabletop's being a model of Euclidean geometry *both* on the customary metric $ds^2 = dx^2 + dy^2$ *and* on the different metric $d\bar{s}^2 = \sec^2\theta \, dx^2 + dy^2$: the geodesics as well as the geometries furnished by these incompatible metrics are the same, but the angles which are congruent in the new metric are generally not congruent in the original one. For these two metrics of the surface are not related by a conformal transformation as defined in Chapter One, Section C, and the conformal relation between the two metrics is a necessary and not only a sufficient condition for the sameness of their associated angle congruences.² That these two metrics issue in *incompatible* congruence classes of *angles* though in the *same* geometry can also be seen very simply as follows: a Euclidean triangle which is equilateral on the new metric $d\bar{s}$ will *not* be equilateral on the customary one ds , and hence the three angles of such a triangle will all be congruent to each other in the former metric but not in the latter.

It is clear now that an *arbitrary* change in the congruence definition for either line segments or angles or both cannot as such guarantee a different geometry.

(2) In reply to Hugo Dingler's contention³ that the rigid body is uniquely specified by the geometry and only by the latter,

² Cf. D. J. Struik: *Differential Geometry*, *op. cit.*, pp. 169–70.

³ H. Dingler: "Die Rolle der Konvention in der Physik," *Physikalische Zeitschrift*, Vol. XXIII (1922), p. 50.

Reichenbach mistakenly agrees⁴ that the geometry is sufficient to define congruence and contests only Dingler's further claim that it is necessary.⁵

Carnap⁶ discusses the dependencies obtaining between (a) the metric geometry, which he symbolizes by "R" in this German publication; (b) the topology of the space and the facts concerning the coincidences of the rod in it, symbolized by "T" for "Tatbestand"; and (c) the metric M ("Mass-setzung"), which entails a congruence definition and is given by the function f (and by the choice of a unit), as will be recalled from the beginning of this chapter.⁷ And he concludes that the functional relations between R, M, and T are such "that if two of them are given, the third specification is thereby uniquely given as well."⁸ Accordingly, he writes:

$$\begin{aligned} R &= \Phi_1 (M, T) \\ M &= \Phi_2 (R, T) \\ T &= \Phi_3 (M, R). \end{aligned}$$

While the first of these dependencies does hold, our example of imparting a Euclidean geometry to a tabletop by each of two incompatible congruence definitions shows that not only the second but also the third of Carnap's dependencies fails to hold. For the mere specification of M to the effect that the rod will be called congruent to itself everywhere and of R as Euclidean does not tell us whether the coincidence behavior T of the rod on the tabletop will be such as to coincide successively with those intervals that are equal according to the formula

⁴ H. Reichenbach: "Discussion of Dingler's Paper," *Physikalische Zeitschrift*, Vol. XXIII (1922), p. 52, and H. Reichenbach: "Über die physikalischen Konsequenzen der relativistischen Axiomatik," *Zeitschrift für Physik* Vol. XXXIV (1925), p. 35.

⁵ In Chapter Four we shall assess the merits of Reichenbach's denial that the geometry is necessary, which he rests on the grounds that rigidity can be defined by the elimination of differential forces.

⁶ R. Carnap: *Der Raum*, *op. cit.*, p. 54.

⁷ Although both Carnap's metric M and the distance function

$$ds = \sqrt{g_{ik}} dx^i dx^k$$

can provide a congruence definition, they cannot be deduced from one another without information concerning the coincidence behavior of the rod in the space under consideration.

⁸ R. Carnap: *Der Raum*, *op. cit.*, p. 54.

$ds^2 = dx^2 + dy^2$ or with the different intervals that are equal on the basis of one of the metrizations $ds^2 = \sec^2\theta \, dx^2 + dy^2$ (where $\sec^2\theta > 1$). In other words, the stated specifications of M and R do not tell us whether the rod behaves on the tabletop as we know it to behave in actuality or whether it squirms in any one of infinitely many different ways as compared to its actual behavior.

(3) As a corollary of our proof of the *non-uniqueness* of the congruence definition, we can show that the following statement by Reichenbach is false: "If we say: actually a geometry G applies but we measure a geometry G', we define at the same time a force F which causes the difference between G and G'."⁹ Using our previous notation, we note first that instead of determining a metric tensor g'_{ik} uniquely (up to an arbitrary constant), the geometry G determines an infinite class α of such tensors differing other than by being proportional to one another. But since $F_{ik} = g_{ik} - g'_{ik}$ (where the g'_{ik} are furnished by the rod prior to its being regarded as "deformed" by any universal forces), the failure of G to determine a tensor g_{ik} uniquely (up to an arbitrary constant) issues in there being as many *different* universal forces F_{ik} as there are different tensors g_{ik} in the class α determined by G. We see, therefore, that contrary to Reichenbach, there are infinitely many different ways in which the measuring rod can be held to be "deformed" while furnishing the *same* geometry G.

These criticisms of Reichenbach do not affect substantial portions of his immensely valuable contributions to the philosophy of geometry. But it is evident that it is less than discerning to draw the following conclusion with E. H. Hutten in regard to the logical analysis of the concept of space as understood since the advent of Einstein's relativity: "except for some change in terminology there is nothing to add to the exposition as found for instance, in Reichenbach's book on the Philosophy of Space and Time."¹

⁹ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 27.

¹ E. H. Hutten: *The Language of Modern Physics* (London: George Allen and Unwin, Ltd., and New York: The Macmillan Company; 1956), p. 110.

Chapter 4

CRITIQUE OF EINSTEIN'S PHILOSOPHY OF GEOMETRY

(A) AN APPRAISAL OF DUHEM'S ACCOUNT OF THE FALSIFIABILITY OF ISOLATED EMPIRICAL HYPOTHESES IN ITS BEARING ON EINSTEIN'S CONCEPTION OF THE INTERDEPENDENCE OF GEOMETRY AND PHYSICS.

Since Einstein's central thesis concerning the epistemological status of physical geometry will be seen in Section C to be a geometrical version of Pierre Duhem's conception of the falsifiability of isolated empirical hypotheses, this first section will be devoted to a critical examination of Duhem's conception as articulated by W. V. O. Quine.

It has been maintained by writers other than Duhem and Quine that there is an important *asymmetry* between the *verification* and the *refutation* of a theory in empirical science. Refutation has been said to be conclusive or decisive while verification was claimed to be irremediably inconclusive in the following sense: If a theory T_1 entails observational consequences O , then the *truth* of T_1 does not, of course, follow deductively from the truth of the conjunction

$$(T_1 \rightarrow O) \cdot O.$$

On the other hand, the *falsity* of T_1 is indeed deductively inferable by *modus tollens* from the truth of the conjunction

$$(T_1 \rightarrow O) \cdot \sim O.$$

Thus, F. S. C. Northrop writes: "We find ourselves, therefore, in this somewhat shocking situation: the method which natural science uses to check the postulationally prescribed theories . . . is absolutely trustworthy when the proposed theory is not confirmed and logically inconclusive when the theory is experimentally confirmed."¹

Under the influence of Duhem,² this thesis of asymmetry of conclusiveness between verification and refutation has been strongly denied as follows: If " T_1 ," denotes the kind of individual or *isolated* hypothesis H whose verification or refutation is at issue in the conduct of particular scientific experiments, then Northrop's formal schema is a misleading oversimplification. Upon taking cognizance of the fact that the observational consequences (O) are deduced *not* from H alone but rather from the conjunction of H and the relevant body of *auxiliary* assumptions A, the refutability of H is seen to be no more conclusive than its verifiability. For now it appears that Northrop's formal schema must be replaced by the following:

$$(i) \quad [(H \cdot A) \rightarrow O] \cdot O \quad (\text{verification})$$

and

$$(ii) \quad [(H \cdot A) \rightarrow O] \cdot \sim O. \quad (\text{refutation}).$$

The recognition of the presence of the auxiliary assumptions A in both the verification and refutation of H now makes apparent that the *refutation* of H *itself* by *adverse* empirical evidence $\sim O$ can be no more decisive than its *verification* (confirmation) by *favorable* evidence O. What can be inferred deductively from the refutational premise (ii) is *not* the falsity of H itself but only the much weaker conclusion that H and A cannot both be true. It is immaterial here that the *falsity* of the *conjunction* of H and A *can* be inferred *deductively* from the refutational premise (ii) while the truth of that conjunction can be inferred only *inductively* from the verificational premise (i). For this does not detract from the fact that there is parity of inconclusiveness between the refutation of H itself and the verification of H itself

¹ F. S. C. Northrop: *The Logic of the Sciences and the Humanities* (New York: The Macmillan Company; 1947), p. 146.

² Pierre Duhem: *The Aim and Structure of Physical Theory* (Princeton: Princeton University Press; 1954), Part II, Ch. vi, especially pp. 183-90.

in the following sense: (ii) does not entail (deductively) the falsity of *H* itself, just as (i) does not entail the truth of *H* by itself. In short, isolated component hypotheses of far flung theoretical systems are not separately refutable but only contextually disconfirmable: *no one constituent hypothesis H* can ever be *extricated* from the ever-present web of collateral assumptions so as to be *open to separate refutation* by the evidence as part of an *explanans* of that evidence, just as no such isolation is achievable for purposes of verification. And Northrop's schema is an adequate representation of the actual logical situation only if "*T*₁" in his schema refers to the entire theoretical *system* of premises which enters into the deduction of *O* rather than to such mere *components H* as are at issue in specific scientific inquiries.

Under the influence of Duhem's emphasis on the confrontation of an entire theoretical system by the tribunal of evidence, writers such as W. V. O. Quine have made what I take to be the following claim: no matter what the specific content *O'* of the *prima facie* adverse empirical evidence $\sim O$, we can always justifiably affirm the truth of *H* as part of the theoretical *explanans* of *O'* by doing two things: First, blame the falsity of *O* on the falsity of *A* rather than on the falsity of *H*, and second, so modify *A* that the conjunction of *H* and the *revised* version *A'* of *A* does entail (explain) the actual findings *O'*. Thus, in his "Two Dogmas of Empiricism," Quine writes: "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system."³ And one of Quine's arguments in that provocative essay against the tenability of the analytic-synthetic distinction is that a supposedly synthetic statement, no less than a supposedly analytic one can be claimed to be true "come what may" on Duhemian grounds. Accordingly, we take the Duhem-Quine thesis—hereafter called the "D-thesis"—to involve the following set of contentions: *there is an inductive (epistemological) interdependence and inseparability between H and the auxiliary assumptions A*, and there is therefore an ingression of a kind of a priori choice into physical theory. For at the price of suitable compensatory modifications in the remainder of the theory, any one of its component hypotheses *H* may be retained in the face

³ W. V. O. Quine: *From a Logical Point of View* (2nd ed.; Cambridge: Harvard University Press; 1961), p. 43. Cf. also p. 41, n. 17.

of seemingly contrary empirical findings as part of an *explanans* of these very findings. And this quasi a priori preservability of H is sanctioned by the far-reaching theoretical ambiguity and flexibility of the logical constraints imposed by the observational evidence.⁴

In the particular case of physical geometry, Duhem would point to the fact that in a sense to be specified *in detail* in Section C, the physical laws used to correct a measuring rod for substance-specific distortions presuppose a geometry and comprise the laws of optics. And hence he would *deny*, for example, that either of the following kinds of *independent* tests of geometry and optics are feasible:

1. Prior to and independently of knowing or presupposing the geometry, we find it to be a law of optics that the paths of light coincide with the geodesics of the congruence defined by rigid bodies.

Knowing this, we then use triangles consisting of a geodesic base line in the solar system and the stellar light rays connecting its extremities to various stars to determine the geometry of the system of rigid body geodesics: stellar parallax measurements will tell us whether the angle sums of the triangles are 180° (Euclidean geometry), less than 180° (hyperbolic geometry) or in excess of 180° (spherical geometry).

If we thus find that the angle sum is different from 180° , then we shall know that the geometry of the rigid body geodesics is not Euclidean. For in view of our prior independent ascertainment of the paths of light rays, such a non-Euclidean result could not be interpreted as due to the failure of optical paths to coincide with the rigid body geodesics.

2. Prior to and independently of knowing or presupposing the laws of optics, we ascertain what the geometry is relatively to the rigid body congruence.

Knowing this we then find out whether the paths of light rays coincide with the geodesics of the rigid body congruence by

⁴ P. Duhem: *The Aim and Structure of Physical Theory*, *op. cit.* Duhem's explicit disavowal of both decisive falsifiability and crucial verifiability of an *explanans* will not bear K. R. Popper's reading of him [K. R. Popper: *The Logic of Scientific Discovery*, *op. cit.*, p. 78]: Popper, who is an exponent of decisive falsifiability [*ibid.*], misinterprets Duhem as allowing that tests of a hypothesis may be decisively falsifying and as denying only that they may be crucially verifying.

making a parallax or some other determination of the angle sum of a light ray triangle.

Since we know the geometry of the rigid body geodesics independently of the optics, we know what the corresponding angle sum of a triangle whose sides are geodesics ought to be. And hence the determination of the angle sum of a light ray triangle is then decisive in regard to whether the paths of light rays coincide with the geodesics of the rigid body congruence.

In place of such independent confirmability and falsifiability of the geometry and the optics, Duhem affirms their *inductive* (epistemological) inseparability and interdependence.

In the present chapter, I shall endeavor, among other things, to establish two main conclusions:

(1) Quine's formulation of Duhem's thesis—which we call the “D-thesis”—is true *only* in various *trivial* senses of what Quine calls “drastic enough adjustments elsewhere in the system.” And no one would wish to contest any of these thoroughly uninteresting versions of the D-thesis,

(2) in its *non-trivial* exciting form, the D-thesis is untenable in the following fundamental respects:

A. *Logically*, it is a *non-sequitur*. For *independently* of the particular empirical context to which the hypothesis H pertains, there is no logical guarantee at all of the existence of the *required* kind of revised set A' of auxiliary assumptions such that

$$(H \cdot A') \rightarrow O'$$

for any one component hypothesis H and any O'. Instead of being guaranteed logically, the existence of the required set A' needs *separate* and *concrete* demonstration for each particular context. In the absence of the latter kind of *empirical* support for Quine's unrestricted Duhemian claim, that claim is an unempirical dogma or article of faith which the pragmatist Quine is no more entitled to espouse than an empiricist would be.

B. The D-thesis is not only a *non-sequitur* but is actually *false*, as shown by an important counter-example, namely the separate falsifiability of a particular component hypothesis H.

Of these conclusions, (1) and (2A) will be defended in the present Section A, while arguments in support of (2B) will be deferred until Section C.

To forestall misunderstanding, let it be noted that my rejection of the very strong assertion made by Quine's D-thesis is not at all intended as a repudiation of the following far weaker contention, which I believe to be eminently sound: the logic of every disconfirmation, no less than of every confirmation of an isolated scientific hypothesis H is such as to *involve at some stage or other* an entire network of interwoven hypotheses in which H is ingredient rather than in every stage merely the separate hypothesis H. Furthermore, it is to be understood that the issue before us is the logical one whether in principle every component H is unrestrictedly preservable by a suitable A', not the psychological one whether scientists possess sufficient ingenuity at every turn to propound the required set A', *if it exists*. Of course, *if* there are cases in which the requisite A' simply *does not even exist logically*, then surely no amount of ingenuity on the part of scientists will enable them to ferret out the non-existent required A' in such cases.

I. *The Trivial Validity of the D-Thesis.*

It can be made evident at once that unless Quine restricts in very specific ways what he understands by "drastic enough adjustments elsewhere in the [theoretical] system," the D-thesis is a thoroughly unenlightening truism. For if someone were to put forward the false empirical hypothesis H that "Ordinary buttermilk is highly toxic to humans," this hypothesis could be saved from refutation in the face of the observed wholesomeness of ordinary buttermilk by making the following "drastic enough" adjustment in our system: changing the rules of English usage so that the intension of the term "ordinary buttermilk" is that of the term "arsenic" in its customary usage. Hence a *necessary* condition for the non-triviality of Duhem's thesis is that *the theoretical language be semantically stable* in the relevant respects.

Furthermore, it is clear that if one were to countenance that O' itself qualifies as A', Duhem's affirmation of the existence of an A' such that

$$(H \cdot A') \rightarrow O'$$

would hold trivially, and H would not even be needed to deduce O'. Moreover, the D-thesis can hold trivially even in cases in

which H is required in addition to A' to deduce the *explanandum* O': an A' of the trivial form

$$\sim H \vee O'$$

requires H for the deduction of O', but no one will find it enlightening to be told that the D-thesis can thus be sustained.

I am unable to give a formal *and* completely general *sufficient* condition for the *non-triviality* of A'. And, so far as I know, neither the originator nor any of the advocates of the D-thesis have even shown any awareness of the need to circumscribe the class of *non-trivial* revised auxiliary hypotheses A' so as to render the D-thesis interesting. I shall therefore assume that the proponents of the D-thesis intend it to stand or fall on the kind of A' which we would all recognize as *non-trivial* in *any given case*, a kind of A' which I shall symbolize by A'_{nt}. And I shall endeavor to show that such a *non-trivial* form of the D-thesis is indeed untenable after first commenting on the attempt to sustain the D-thesis by resorting to the use of a *non-standard logic*.

The species of drastic adjustment consisting in recourse to a *non-standard logic* is specifically mentioned by Quine. Citing a hypothesis such as "there are brick houses on Elm Street," he claims that even a statement so "germane to sense experience . . . can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws."⁵ I disregard for now the argument from hallucination. In the absence of *specifics* as to the ways in which alterations of logical laws will enable Quine to hold in the face of *recalcitrant experience* that a statement H like "there are brick houses on Elm Street" is *true*, I must conclude the following: the invocation of *non-standard logics* either makes the D-thesis *trivially* true or turns it into an interesting claim which is an unfounded *dogma*. For suppose that the *non-standard logic* used is a *three-valued* one. Then even if it were otherwise feasible to assert within the framework of such a logic that the particular statement H is "true," the term "true" would no longer have the meaning associated with the *two-valued* framework of logic within which the D-thesis was enunciated to begin with. It is not to be overlooked that a form of the D-thesis which allows it-

⁵ W. V. O. Quine: *From a Logical Point of View*, *op. cit.*, p. 43.

self to be sustained by alterations in the meaning of "true" is no less trivial *in the context of the expectations raised by the D-thesis* than one which rests its case on calling arsenic "butter-milk." And this triviality obtains *in this context*, notwithstanding the fact that the two-valued and three-valued usages of the word "true" share what H. Putnam has usefully termed a common "core meaning."⁶

For suppose we had two particular substances I_1 and I_2 which are isomeric with each other. That is to say, these substances are composed of the same elements in the same proportions and with the same molecular weight but the arrangement of the atoms within the molecule is different. Suppose further that I_1 is not at all toxic while I_2 is highly toxic, as in the case of two isomers of trinitrobenzene.⁷ Then if we were to call I_1 "duquine" and asserted that "duquine is highly toxic," this statement H could also be trivially saved from refutation in the face of the evidence of the wholesomeness of I_1 by the following device: only *partially* changing the meaning of "duquine" so that its intension is the second, highly toxic isomer I_2 , thereby leaving the chemical "core meaning" of "duquine" intact. To forestall the *misunderstanding* that my charge of triviality here is guilty of the error committed by the Eddington-Putnam triviality-thesis (cf. Chapter One, Section D), let me point out precisely what I regard as trivial in this context. The preservation of H from refutation in the face of the evidence by a *partial* change in the meaning of "duquine" is trivial in the sense of being only a *trivial* fulfillment of the expectations raised by the D-thesis. But, in my view, the possibility as such of preserving H *by this particular kind of change in meaning* is not at all trivial. For this possibility as such reflects a *fact about the world*: the existence of isomeric substances of radically different degrees of toxicity (allergenicity)!⁸

Even if one ignores the change in the meaning of "true"

⁶ H. Putnam: "Three-Valued Logic," *Philosophical Studies*, Vol. VIII (1957), p. 74.

⁷ Cf. H. L. Alexander: *Reactions With Drug Therapy* (Philadelphia: W. B. Saunders Co.; 1955), p. 14. Alexander writes: "It is true that drugs with closely related chemical structures do not always behave clinically in a similar manner, for antigenicity of simple chemical compounds may be changed by minor alterations of molecular structures. . . . 1,2,4-trinitrobenzene . . . is a highly antigenic compound. . . . 1,3,5 . . . trinitrobenzene is allergenically inert." (I am indebted to Dr. A. I. Braude for this reference.)

inherent in the resort to a three-valued logic, there is no reason to think that the D-thesis can be successfully upheld in such an altered logical framework: the arguments which I shall present against the *non*-trivial form of the D-thesis within the framework of the standard logic apply just as much, so far as I can see, in the three-valued and other non-standard logics of which I am aware. And if the reply is that there are other non-standard logics which are both viable for the purposes of science and in which my impending polemic against the non-trivial form of the D-thesis does not apply, then I retort: as it stands, Quine's assertion of the feasibility of a change in the laws of logic which would thus sustain the D-thesis is an unempirical dogma or at best a promissory note. And until the requisite collateral is supplied, it is not incumbent upon anyone to accept that promissory note.

II. *The Untenability of the Non-Trivial D-Thesis.*

The *non*-trivial D-thesis will now be seen to be a *non-sequitur*. The non-trivial D-thesis is that for every component hypothesis H of any domain of empirical knowledge and for any observational findings O',

$$(\exists A'_{nt}) [(H \cdot A'_{nt}) \rightarrow O'].$$

But this claim does not follow from the fact that the falsity of H is not deductively inferable from the premise

$$[(H \cdot A) \rightarrow O] \cdot \sim O,$$

which we shall call premise (ii) as at the beginning of this section. For the latter premise utilizes not the full empirical information given by O' but only the part of that information which tells us that O' is logically incompatible with O. Hence the *failure* of $\sim O$ to permit the deduction of $\sim H$ does *not* justify the assertion of the D-thesis that there always *exists* a non-trivial A' such that the conjunction of H and that A' entails O'. In other words, the fact that the falsity of H is *not* deducible (by *modus tollens*) from premise (ii) is quite insufficient to show that H can be preserved non-trivially as part of an *explanans* of *any* potential empirical findings O'. I conclude, therefore, from the analysis given so far that in its *non*-trivial form,

Quine's D-thesis is *gratuitous* and that the existence of the required non-trivial A' would require *separate* demonstration for each particular case.

(B) THE INTERDEPENDENCE OF GEOMETRY AND PHYSICS
IN POINCARÉ'S CONVENTIONALISM.

The literature treating of the interdependence of geometry and physics exhibits a pervasive confusion between two logically very different kinds of interdependence which can be respectively associated with Duhem's epistemological holism and Poincaré's conventionalism: an *inductive* (*epistemological*) interdependence and a *linguistic* one. The distinction between these two kinds of interdependence merits being drawn in its own right and is also of considerable importance for the critical estimate of Einstein's conception to be given in Section C.

In the present section, I shall therefore discuss the relevant difference between the views of Duhem and Poincaré as well as some related questions pertaining to the interpretation of Poincaré.

As we saw in Section C of Chapter One, the central theme of Poincaré's so-called conventionalism is essentially an elaboration of the thesis of alternative metrizableability whose fundamental justification we owe to Riemann. Poincaré's much cited but widely misunderstood statement concerning the possibility of always giving a Euclidean description of any results of stellar parallax measurements reads as follows:

If Lobachevski's geometry is true, the parallax of a very distant star will be finite; if Riemann's is true, it will be negative. These are results which seem within the reach of experiment, and there have been hopes that astronomical observations might enable us to decide between the three geometries.

But in astronomy "straight line" means simply "path of a ray of light."

If therefore negative parallaxes were found, or if it were demonstrated that all parallaxes are superior to a certain limit, two courses would be open to us; we might either renounce Euclidean geometry, or else modify the laws of optics and suppose that light does not travel rigorously in a straight line.

It is needless to add that all the world would regard the latter solution as the more advantageous.

The Euclidean geometry has, therefore, nothing to fear from fresh experiments.⁸

The context of this paragraph⁹ makes it quite clear that Poincaré is using the observationally significant case of a stellar light ray triangle to explain that, if need be, the preservation of a Euclidean description by an alternative metrization is a genuinely live option. Hence, his account of the relevance of stellar parallax measurements to the determination of the metric geometry of physical space makes precisely the same point, albeit much less lucidly, as the following magisterially clear statement by him:

In space we know rectilinear triangles the sum of whose angles is equal to two right angles; but equally we know curvilinear triangles the sum of whose angles is less than two right angles. . . . To give the name of straights to the sides of the first is to adopt Euclidean geometry; to give the name of straights to the sides of the latter is to adopt the non-Euclidean geometry. So that to ask what geometry it is proper to adopt is to ask, to what line is it proper to give the name straight? It is evident that experiment can not settle such a question.¹

Now, the equivalence of this latter contention to Riemann's view of congruence becomes evident the moment we note that the legitimacy of identifying lines which are curvilinear in the usual geometrical parlance as "straights" is vouchsafed by the warrant for our choosing a new definition of congruence such that the previously curvilinear lines become geodesics of the new congruence.

And we note that whereas the *original* geodesics in space exemplified the formal relations obtaining between Euclidean "straight lines," the *different* geodesics associated with the new metrization embody the relations prescribed for straight lines by the formal postulates of *hyperbolic* geometry. Awareness of the fact that Poincaré begins the quoted passage with the words "In [physical] space" enables us to see that he is making the follow-

⁸ H. Poincaré: *The Foundations of Science*, *op. cit.*, p. 81.

⁹ *Ibid.*, pp. 81-86.

¹ *Ibid.*, p. 235.

ing assertion here: the same *physical* surface or region of three-dimensional *physical* space admits of *alternative* metrizations so as to constitute a physical realization of either the formal postulates of Euclidean geometry or of one of the non-Euclidean abstract calculi. To be sure, syntactically, this alternative metrization involves a formal intertranslatability of the *relevant portions* of these incompatible geometrical calculi, the "intertranslatability" being guaranteed by a "dictionary" which pairs off with one another the *alternative names* (or descriptions) of each physical path or configuration. But the essential point made here by Poincaré is not that a purely formal translatability obtains; instead, Poincaré is emphasizing here that a given physical surface or region of physical 3-space can indeed be a model of one of the *non-Euclidean* geometrical calculi no less than of the Euclidean one. In this sense one can say, therefore, that Poincaré affirmed the conventional or definitional status of applied geometry.

Hence, we must reject the following wholly syntactical interpretation of the above citation from Poincaré, which is offered by Ernest Nagel, who writes: "The thesis he [Poincaré] establishes by this argument is simply the thesis that choice of notation in formulating a system of pure geometry is a convention."² Having thus misinterpreted Poincaré's conventionalist thesis as pertaining only to formal intertranslatability, Nagel fails to see that Poincaré's avowal of the conventionality of physical or *applied* geometry is none other than the assertion of the *alternative metrization of physical space* (or of a portion thereof). And, in this way, Nagel is driven to give the following unfounded interpretation of Poincaré's conception of the status of applied (physical) geometry: "Poincaré also argued for the definitional status of *applied* as well as of *pure* geometry. He maintained that, even when an interpretation is given to the primitive terms of a pure geometry so that the system is then converted into statements about certain physical configurations (for example, interpreting 'straight line' to signify the path of a light ray), no experiment on physical geometry can ever decide against one of the alternative systems of physical geometry and in favor of another."³ But far

² E. Nagel: *The Structure of Science*, *op. cit.*, p. 261.

³ *Ibid.*

from having claimed that the geometry is still conventional even *after* the provision of a particular physical interpretation of a pure geometry, Poincaré merely reiterated the following thesis of alternative metrizable in the passages which Nagel⁴ then goes on to quote from him: suitable alternative semantical interpretations of the term "congruent" (for line segments and or for angles), and correlatively of "straight line," etc., can readily demonstrate that, subject to the restrictions imposed by the existing topology, it is always a live option to give *either* a Euclidean *or* a *non*-Euclidean description of any given set of *physico*-geometric facts. And since alternative metrizations are just as legitimate epistemologically as alternative systems of units of length or temperature, one can always, in principle, *reformulate* any physical theory based on a given metrization of space—or, as we saw in Chapter Two above, of time—so as to be based on an *alternative* metric.

There is therefore no warrant at all for the following caution expressed by Nagel in regard to the feasibility of what is merely a reformulation of physical theory on the basis of a new metrization: "... even if we admit universal forces in order to retain Euclid... we must incorporate the assumption of universal forces into the rest of our physical theory, rather than introduce such forces piecemeal subsequent to each observed 'deformation' in bodies. It is by no means self-evident, however, that physical theories can in fact always be devised that have built-in provisions for such universal forces."⁵ Yet, precisely that fact is self-evident, and its self-evidence is obscured from view by the logical havoc created by the statement of a remetrization issuing in Euclidean geometry in terms of "universal forces." For that metaphor seems to have misled Nagel into imputing the status of an *empirical* hypothesis to the use of a *non*-standard spatial metric merely because the latter metric is described by saying that we "assume" appropriate universal forces. In fact, our discussion in Chapter Two has shown mathematically for the one-dimensional case of time how Newtonian mechanics is to be recast via suitable transformation equations, such as equation (4) there, so as to implement a remetrization given by $T = f(t)$, which can be

⁴ *Ibid.*, pp. 261–62.

⁵ *Ibid.*, pp. 264–65.

described metaphorically by saying that all clocks are "accelerated" by "universal forces."

Corresponding remarks apply to Poincaré's contention that we can always preserve Euclidean geometry in the face of any data obtained from stellar parallax measurements: if the paths of light rays are geodesics on the customary definition of congruence, as indeed they are in the Schwarzschild procedure cited by Robertson,⁶ and if the paths of light rays are found parallaxically to sustain non-Euclidean relations on that metrization, then we need only choose a different definition of congruence such that these same paths will no longer be geodesics and that the geodesics of the newly chosen congruence are Euclideanly related. From the standpoint of synthetic geometry, the latter choice effects a *renaming* of optical and other paths and thus is merely a *recasting of the same factual content in Euclidean language rather than a revision of the extra-linguistic content of optical and other laws*. The retainability of Euclideanism by remetrization, which is affirmed by Poincaré, therefore *involves a merely linguistic interdependence of the geometric theory of rigid solids and the optical theory of light rays*. And since Poincaré's claim here is a straightforward elaboration of the metric amorphousness of the continuous manifold of space, it is not clear how H. P. Robertson⁷ can reject it as a "pontifical pronouncement" and even regard it as being in contrast with what he calls Schwarzschild's "sound operational approach to the problem of physical geometry." For Schwarzschild had rendered the question concerning the prevailing geometry *factual* only by the adoption of a particular spatial metrization based on the travel times of light, which does indeed turn the direct light paths of his astronomical triangle into geodesics.⁸

See Append §6

⁶ H. P. Robertson: "Geometry as a Branch of Physics," *Albert Einstein: Philosopher-Scientist*, *op. cit.*, pp. 324–25.

⁷ *Ibid.*, pp. 324–25.

⁸ For very useful discussions of the actual astronomical methods used to determine the geometry of physical space in the large, see H. P. Robertson, *op. cit.*, pp. 323–25 and 330–32; Max Jammer's historical work *Concepts of Space* (Cambridge: Harvard University Press; 1954), pp. 147–48; William A. Baum: "Photoelectric Test of World Models," *Science*, Vol. CXXXIV (1961), p. 1426, and Allan Sandage: "Travel Time for Light from Distant Galaxies Related to the Riemannian Curvature of the Universe," *Science*, Vol. CXXXIV (1961), p. 1434.

Poincaré's interpretation of the parallactic determination of the geometry of a stellar triangle has also been obscured by Ernest Nagel's statement of it. Apart from encumbering that statement with the metaphorical use of "universal forces," Nagel fails to point out that the crux of the preservability (retainability) of Euclidean geometry lies in: First, the denial of the *geodesicity* (*straightness*) of optical paths which are found parallactically to sustain *non*-Euclidean relations on the *customary* metrization of line segments and angles, or at least in the rejection of the customary congruence for *angles* (cf. Chapter Three, Section B),⁹ and second, the ability to guarantee the existence of a suitable new metrization whose associated geodesics are paths which do exhibit the formal relations of Euclidean straights. For Nagel characterizes the retainability of Euclidean geometry come what may by asserting that the latter's retention is effected "only by maintaining that the sides of the stellar triangles are not really Euclidean [*sic*] straight lines, and he [the Euclidean geometer] will therefore adopt the hypothesis that the optical paths are deformed by some fields of force."¹ But apart from the obscurity of the notion of the deformation of the optical *paths*, the unfortunate inclusion of the word "Euclidean" in this sentence of Nagel's obscures the very point which the advocate of Euclid is concerned to make in this context in the interests of his thesis. And this point is not, as Nagel would have it, that the optical paths are not really *Euclidean* straight lines, a fact whose admission (assuming the customary congruence for angles) provided the starting point of the discussion. Instead, what the proponent of Euclid is concerned to point out here is that the legitimacy of alternative metrizations enables him to offer a metric such that the optical paths do *not* qualify as *geodesics* (straights) from the

⁹ Under the assumed conditions as to the parallactic findings, a new metrization may allow the optical paths to be interpretable as *Euclideanly*-related geodesics, but only if the customary *angle* congruences were abandoned and changed suitably as part of the remetrization (cf. Chapter Three, Section B). In that case, the paths of light rays would be straight lines *even in the Euclidean description* obtained by the new metrization, but the optical laws involving angles would have to be suitably restated. For theorems governing the so-called "geodesic mapping" or "geodesic correspondence" relevant here, cf. L. P. Eisenhart: *An Introduction to Differential Geometry*, *op. cit.*, Sec. 37, pp. 205-11, and D. J. Struik: *Differential Geometry*, *op. cit.*, pp. 177-80.

¹ E. Nagel: *The Structure of Science*, *op. cit.*, p. 263.

outset. For it is by denying altogether the *geodesicity* of the optical paths that the advocate of Euclid can uphold his thesis successfully in the face of the admitted *prima facie non-Euclidean* parallactic findings.

The invocation of the conventionality of congruence to carry out remetrizations is not at all peculiar to Poincaré. For F. Klein's relative consistency proof of hyperbolic geometry via a model furnished by the interior of a circle on the Euclidean plane,² for example, is based on one particular kind of possible remetrization of the circular portion of that plane, projective geometry having played the heuristic role of furnishing Klein with a suitable definition of congruence. Thus what from the point of view of synthetic geometry appears as intertranslatability via a dictionary, appears as alternative metrizability from the point of view of differential geometry.

There are two respects, however, in which Poincaré is open to criticism in this connection:

First, he maintained³ that it would always be regarded as most convenient to preserve Euclidean geometry, even at the price of remetrization, on the grounds that this geometry is the simplest analytically.⁴ As is well-known, precisely the opposite development materialized in the GTR: Einstein forsook the simplicity of the geometry itself in the interests of being able to maximize the simplicity of the definition of congruence. He makes clear in his fundamental paper of 1916 that had he insisted on the retention of Euclidean geometry in a gravitational field, then he could *not* have taken "one and the same rod, independently of its place and orientation, as a realization of the same interval."⁵

Second, even if the simplicity of the geometry itself were the sole determinant of its adoption, that simplicity might be judged by criteria other than Poincaré's analytical simplicity such as the simplicity of the undefined concepts used.

However, if Poincaré were alive today, he could point to an interesting recent illustration of the sacrifice of the simplicity and accessibility of the congruence standard on the altar of maximum

² R. Bonola: *Non-Euclidean Geometry*, *op. cit.*, pp. 164–75.

³ H. Poincaré: *The Foundations of Science*, *op. cit.*, p. 81.

⁴ *Ibid.*, p. 65.

⁵ A. Einstein: "The Foundations of the General Theory of Relativity," in *The Principle of Relativity*, *op. cit.*, p. 161.

simplicity of the resulting theory. Astronomers have recently proposed to remetrize the time continuum for the following reason: as indicated in Chapter Two, when the mean solar second, which is a very precisely known fraction of the period of the earth's rotation on its axis, is used as a standard of temporal congruence, then there are three kinds of discrepancies between the actual observational findings and those predicted by the usual theory of celestial mechanics. The empirical facts thus present astronomers with the following choice: either they retain the rather natural standard of temporal congruence at the cost of having to bring the principles of celestial mechanics into conformity with observed fact by revising them appropriately, or they remetrize the time continuum, employing a less simple definition of congruence so as to preserve these principles intact. Decisions taken by astronomers in the last few years were exactly the reverse of Einstein's choice of 1916 as between the simplicity of the standard of congruence and that of the resulting theory. The mean solar second is to be supplanted by a unit to which it is non-linearly related: the sidereal year, which is the period of the earth's revolution around the sun, due account being taken of the irregularities produced by the gravitational influence of the other planets.⁶

We see that the implementation of the requirement of descriptive simplicity in theory-construction can take alternative forms, because agreement of astronomical theory with the evidence now available is achievable by revising either the definition of temporal congruence or the postulates of celestial mechanics. The existence of this alternative likewise illustrates that for an axiomatized physical theory containing a geo-chronometry, it is *gratuitous* to single out the postulates of the theory as having been prompted by *empirical* findings in contradistinction to deeming the *definitions of congruence* to be wholly a priori, or vice versa. This conclusion bears out geo-chronometrically Braithwaite's contention⁷ that there is an important sense in which axiomatized physical theory does not lend itself to compliance with Heinrich

⁶ G. M. Clemence: "Time and its Measurement," *op. cit.*

⁷ R. B. Braithwaite: "Axiomatizing a Scientific System by Axioms in the Form of Identification," in *The Axiomatic Method* ed. by L. Henkin, P. Suppes, and A. Tarski (Amsterdam: North Holland Publishing Company; 1959), pp. 429-42.

Hertz's injunction to "distinguish thoroughly and sharply between the elements . . . which arise from the necessities of thought, from experience, and from arbitrary choice."⁸ The same point is illustrated by the possibility of characterizing the factual innovation wrought by Einstein's abandonment of Euclidean geometry in favor of Riemannian geometry in the GTR in several ways as follows:

(1) *Upon using the customary definition of spatial congruence*, the geometry near the sun is *not* Euclidean, contrary to the claims of pre-GTR physics.

(2) The geometry near the sun is *not* Euclidean on the basis of the *customary* congruence, but it *is* Euclidean on a suitably modified congruence definition which makes the length of a rod a specified function of its position and orientation.⁹

(3) *within the confines* of the requirement of giving a *Euclidean* description of the non-classical facts postulated by the GTR, Einstein recognized the *factually dictated* need to *abandon* the *customary* definition of congruence, which had yielded a Euclidean description of the classically assumed facts. Thus, the revision of the Newtonian theory made necessary by the discovery of relativity can be formulated as either a change in the postulates of geometric theory or a change in the correspondence rule for congruence.

Having seen that Poincaré's remetrizational retainability of Euclidean geometry or of some other particular geometry involves a merely *linguistic* interdependence of the geometric theory of rigid solids and the optical theory of light rays, we are ready to contrast that interdependence with the quite different epistemological (inductive) interdependence affirmed by Duhem.

The Duhemian conception envisions scope for alternative geometric accounts of a given body of evidence to the extent that these geometries are associated with alternative, factually *non-equivalent* sets of physical laws which are used to compute *corrections* for substance-specific distortions.¹ Poincaré,² however, in-

⁸ H. Hertz: *The Principles of Mechanics* (New York: Dover Publications, Inc.; 1956), p. 8.

⁹ The function in question is given in R. Carnap: *Der Raum*, *op. cit.*, p. 58.

¹ For some details on just how factually non-equivalent correctional physical laws are associated with different metric geometries, see Section C of this chapter.

² H. Poincaré: *The Foundations of Science*, *op. cit.*, pp. 66–80.

stead of invoking the Duhemian *inductive* latitude, specifically bases the possibility of giving either a Euclidean or a non-Euclidean description of the same spatio-physical facts on alternative metrization *quite apart from any considerations of substance-specific distorting influences and even after correcting for these in some way or other*. Says he: "No doubt, in our world, natural solids . . . undergo variations of form and volume due to warming or cooling. But we neglect these variations in laying the foundations of geometry, because, besides their being very slight, they are irregular and consequently seem to us accidental."³ For the sake of specificity, our contrasting comparison of Poincaré and Duhem will focus on the feasibility of alternative geometric interpretations of stellar parallax data.

See Append. §7

The attempt to explain parallactic data yielding an angle sum different from 180° for a stellar light ray triangle by different geometries which constitute live options *in the inductive sense of Duhem* would presumably issue in the following alternative between two theoretical systems. Each of these theoretical systems comprises a geometry G and an optics O which are epistemologically inseparable and which are inductively interdependent in the sense that the combination of G and O must yield the observed results:

(a) G_E : the geometry of the rigid body geodesics is Euclidean, and

O_1 : the paths of light rays do *not* coincide with these geodesics but form a non-Euclidean system,

or

(b) G_{non-E} : the geodesics of the rigid body congruence are *not* a Euclidean system, and

O_2 : the paths of light rays *do* coincide with these geodesics, and thus they form a non-Euclidean system.

To contrast this Duhemian conception of the feasibility of alternative geometric interpretations of the assumed parallactic data with that of Poincaré, we recall that the physically interpreted alternative geometries associated with two (or more) different metrizations *in the sense of Poincaré* have precisely the same

³ *Ibid.*, p. 76.

total factual content, as do the corresponding two sets of optical laws. For an alternative metrization in the sense of Poincaré affects only the *language* in which the facts of optics and the coincidence behavior of a transported rod are described: the two geometric descriptions respectively associated with two alternative metrizations are *alternative representations of the same factual content*, and so are the two sets of optical laws corresponding to these geometries. Hence I maintain that Poincaré is affirming a *linguistic* interdependence of the geometric theory of rigid solids and the optical theory of light rays. By contrast, in the Duhemian account, G_E and G_{non-E} not only differ in factual content but are logically incompatible, and so are O_1 and O_2 . And on the latter conception, there is sameness of factual content *in regard to the assumed parallactic data* only between the *combined* systems formed by the two conjunctions (G_E and O_1) and (G_{non-E} and O_2).⁴ Thus, the need for the combined system of G and O to yield the empirical facts, coupled with the avowed epistemological (inductive) inseparability of G and O lead the Duhemian to conceive of the interdependence of geometry and optics as inductive (epistemological).

Hence whereas Duhem construes the interdependence of G and O inductively such that the geometry by itself is not accessible to empirical test, Poincaré's conception of their interdependence allows for an empirical determination of G by itself, if we have renounced recourse to an alternative metrization in which the length of the rod is held to vary with its position or orientation. This is not, of course, to say that Duhem regarded alternative metrizations as such to be illegitimate.

It would seem that it was Poincaré's discussion of the interdependence of optics and geometry by reference to stellar parallax measurements which has misled many writers such as Einstein,⁵ Eddington,⁶ and Nagel⁷ into regarding him as a proponent of the Duhemian thesis. An illustration of the widespread conflation of

⁴ These combined systems do not, however, have the same over-all factual content.

⁵ A. Einstein: *Geometrie und Erfahrung* (Berlin: Julius Springer; 1921), p. 9, and A. Einstein: "Reply to Criticisms," in P. A. Schilpp (ed.) *Albert Einstein: Philosopher-Scientist* (Evanston: The Library of Living Philosophers; 1949), pp. 665–88.

⁶ A. S. Eddington: *Space, Time and Gravitation*, *op. cit.*, p. 9.

⁷ E. Nagel: *The Structure of Science*, *op. cit.*, p. 262.

the linguistic and inductive kinds of interdependence of geometry and physics (optics) is given by D. M. Y. Sommerville's discussion of what he calls "the inextricable entanglement of space and matter." He says:

A . . . "vicious circle" . . . arises in connection with the astronomical attempts to determine the nature of space. These experiments are based upon the received laws of astronomy and optics, which are themselves based upon the euclidean assumption. It might well happen, then, that a discrepancy observed in the sum of the angles of a triangle could admit of an explanation by some modification of these laws, or that even the absence of any such discrepancy might still be compatible with the assumptions of non-euclidean geometry.

Sommerville then quotes the following assertion by C. D. Broad:

All measurement involves both physical and geometrical assumptions, and the two things, space and matter, are not given separately, but analysed out of a common experience. Subject to the general condition that space is to be changeless and matter to move about in space, we can explain the same observed results in many different ways by making compensatory changes in the qualities that we assign to space and the qualities we assign to matter. Hence it seems theoretically impossible to decide by any experiment what are the qualities of one of them in distinction from the other.

And Sommerville's immediate comment on Broad's statement is the following:

It was on such grounds that Poincaré maintained the essential impropriety of the question, "Which is the true geometry?" In his view it is merely a matter of convenience. Facts are and always will be most simply described on the euclidean hypothesis, but they can still be described on the non-euclidean hypothesis, with suitable modifications of the physical laws. To ask which is the true geometry is then just as unmeaning as to ask whether the old or the metric system is the true one.⁸

Having dealt with the misinterpretation of Poincaré as a Duhemian, it remains to remove the misunderstanding that Poincaré

⁸ D. M. Y. Sommerville: *The Elements of Non-Euclidean Geometry* (New York: Dover Publications, Inc.; 1958), pp. 209–10. Reprinted through permission of the publisher.

intended his conventionalism as a *denial* of the following Carnap-Reichenbach thesis, which was discussed in Chapter Three: in principle, the question as to the geometry of physical space is *empirical* after the geometrical vocabulary (including the term "congruent" for intervals and angles) has been given a physical interpretation.

According to a widely accepted reading of Poincaré's writings, he is said to have maintained that even after a system of abstract geometry is given a semantical interpretation via a particular coordinative definition of congruence, no experiment can verify or falsify the resulting system of physical geometry, the choice of the particular metrical geometry being entirely a matter of convention."

The principal basis for the belief that Poincaré took a stand in opposition to the kind of qualified metrical empiricism which is espoused by writers such as Reichenbach and Carnap seems to be Poincaré's treatment of "Experience and Geometry" in Chapter Five of his *Science and Hypothesis*,¹ the fifth section of which culminates in the statement that "whichever way we look at it, it is impossible to discover in geometric empiricism a rational meaning."² But there seems to be general unawareness of the fact that Poincaré lifted Sections 4 and 5 of this chapter verbatim out of the wider context of his earlier paper "Des Fondements de la Géométrie, à propos d'un Livre de M. Russell,"³ which was fol-

¹ The view that Poincaré was an extreme geometric conventionalist, who rejected the *qualified* geometric empiricism of Carnap and Reichenbach, is expressed, for example, by the following authors: H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, p. 36; *The Rise of Scientific Philosophy*, *op. cit.*, p. 133, and "The Philosophical Significance of the Theory of Relativity" in: P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist* (Evanston: Library of Living Philosophers; 1949), p. 297; E. Nagel: *The Structure of Science*, *op. cit.*, p. 261; "Einstein's Philosophy of Science," *The Kenyon Review*, Vol. XII (1950), p. 525, and "The Formation of Modern Conceptions of Formal Logic in the Development of Geometry," *Osiris*, Vol. VII (1939), pp. 212-16; H. Weyl: *Philosophy of Mathematics and Natural Science* (Princeton: Princeton University Press; 1949), p. 134, and O. Holder: *Die Mathematische Methode* (Berlin: Julius Springer; 1924), p. 400, n. 2.

² Cf. H. Poincaré: *The Foundations of Science*, *op. cit.*, pp. 81-86.

³ *Ibid.*, p. 86.

⁴ This critique of Russell's *Foundations of Geometry* appeared in *Revue de Métaphysique et de Morale*, Vol. VII (1899), pp. 251-79; the transplanted excerpt is given in Sec. 12, pp. 265-67 of this paper.

lowed by his important rejoinder "Sur les Principes de la Géométrie, Réponse à M. Russell."⁴ These neglected papers together with his posthumous *Dernières Pensées*⁵ seem to me to show convincingly that Poincaré was not an opponent of the qualified kind of empiricist position taken by Reichenbach and Carnap. And I explain his apparent endorsement of unmitigated, anti-empiricist conventionalism in his more publicized writings on the basis of the historical context in which he wrote.

For at the turn of the century, the Riemannian kind of qualified empiricist conception of physical geometry, which takes full cognizance of the stipulational status of congruence and which we now associate with writers like Carnap and Reichenbach, had hardly secured a sufficient philosophical following to provide a stimulus and furnish a target for Poincaré's polemic. Instead, the then dominant philosophical interpretations of geometry were such aprioristic neo-Kantian ones as Couturat's and Russell's, on the one hand, and Helmholtz's type of empiricist interpretation, which made *inadequate* allowance for the stipulational character of congruence, on the other.⁶ No wonder, therefore, that Poin-

⁴ *Revue de Métaphysique et de Morale*, Vol. VIII (1900), pp. 73-86; the relevant paper by Russell is "Sur les Axiomes de la Géométrie," in Vol. VII (1899), pp. 684-707 of that same journal.

⁵ H. Poincaré: *Dernières Pensées* (Paris: Flammarion; 1913), Chaps. 2 and 3.

⁶ Cf. H. von Helmholtz: *Schriften zur Erkenntnistheorie*, ed. by P. Hertz and M. Schlick (Berlin: Julius Springer; 1921), pp. 15-20. H. Freudenthal has maintained (*Mathematical Reviews*, Vol. XXII [1961], p. 107) that instead of being a supporter of Riemann, against Helmholtz, Poincaré was an exponent of Helmholtz's anti-Riemannian view that metric geometry presupposes a three-dimensional rather than a merely one-dimensional solid body as a congruence standard. Freudenthal backs that interpretation of Poincaré by the latter's declaration that "if then there were no solid bodies in nature, there would be no geometry" ("L'Espace et la Géométrie," *Revue de Métaphysique et de Morale*, Vol. III [1895], p. 638). According to Freudenthal ("Zur Geschichte der Grundlagen der Geometrie," *Nieuw Archief voor Wiskunde*, Vol. V [1957], p. 115), this declaration shows that "Poincaré still thinks quite in the empiricist spirit of Helmholtz's space problem and has not even penetrated to Riemann's conception, which is aware of a metric without rigid bodies." But, contrary to Freudenthal, it seems clear from the context of Poincaré's declaration that his mention of the role of solid bodies pertains not at all to a Helmholtzian insistence on a three-dimensional congruence standard as against Riemann's one-dimensional one; instead it concerns the role of solids in the genesis of the notion of mere change of position as against other changes of state, solids being distinguished from liquids and gases by the fact that their displace-

caré's conventionalist emphasis in his better known but incomplete writings seems, in the contemporary context, to place him into the ranks of such extreme conventionalists as H. Dingler.⁷

As evidence for my non-standard interpretation of Poincaré as being a *qualified* geometric empiricist rather than an extreme conventionalist, I cite the following crucial and unequivocal concluding passage from Poincaré's rejoinder to Russell, who had maintained that the "axiom of free mobility" furnishes a uniquely true criterion of congruence as an a priori condition for the possibility of metric geometry in the Kantian presuppositional sense and not in the sense of a coordinative definition. Poincaré writes:

Finally, I have never said that one can *ascertain by experiment whether certain bodies preserve their form*. I have said just the contrary. The term "to preserve one's form" has no meaning by itself. But I confer a meaning on it by *stipulating* that certain bodies will be said to preserve their form. These bodies, thus chosen, can henceforth serve as instruments of measurement. But if I say that these bodies preserve their form, it is because *I choose to do so* and not because experience obliges me to do so.

In the present context I choose to do so, because by a series of *observations* ("constatations") analogous to those which were under discussion in the previous section [i.e., observations showing the coincidence of certain points with others in the course of the movements of bodies] *experience has proven* to me that their movements form a Euclidean group. I have been able to make these observations in the manner just indicated

ments lend themselves to compensation by a corresponding movement of our own bodies, which issues in the restoration of the set of sense impressions we had of the solids prior to their displacement.

But this view is, of course, entirely consonant both with Riemann's conception of the congruence standard as one-dimensional and with his claim that, being continuous, physical space has no intrinsic metric, the latter having to be brought in from elsewhere, as is done by the use of the rigid body. In fact, how except by embracing precisely this view could Poincaré have espoused the conventionality of congruence and the resulting alternative metrizable of physical space on which he founded his thesis of the feasibility of either a Euclidean or a non-Euclidean description?

⁷ Poincaré himself deplored the widespread misunderstandings of his philosophical work and its misappropriation by "all the reactionary French journals." Cf. his *La Mécanique Nouvelle*, cited in R. Dugas: "Henri Poincaré devant les Principes de la Mécanique," *Revue Scientifique*, Vol. LXXXIX (1951), p. 81.

without having any preconceived idea concerning metric geometry. And, having made them, I judge that the convention will be convenient and I adopt it.⁸

It must also be remembered that Poincaré's declaration that "no geometry is either true or false"⁹ was made by him as part of a discussion in which he contrasted his endorsement of this proposition with his complete rejection of the following two others: First, the truth of Euclidean geometry is known to us a priori independently of all experience, and second, one of the geometries is true and the others false, but we can never know which one is true. The entire tenor of this discussion makes it clear that Poincaré is concerned there with abstract, uninterpreted geometries whose relations to physical facts are as yet indeterminate by virtue of the absence of coordinative definitions. It is because he is directing his critique against those who fail to grasp that the identification of the equality predicate "congruent" with its denotata is not a matter of factual truth but of coordinative definition that he asks in *Science and Hypothesis*: "how shall one know [without circularity] that any concrete magnitude which I have measured with my material instrument really represents the abstract distance?"¹

But one might either contest my interpretation here or conclude that Poincaré was inconsistent by pointing to the following passage by him:

Should we . . . conclude that the axioms of geometry are experimental verities? . . . If geometry were an experimental science, it would not be an exact science, it would be subject to a continual revision. Nay, it would from this very day be convicted of error, since we know that there is no rigorously rigid solid.

The axioms of geometry therefore are . . . conventions . . .
Thus it is that the postulates can remain rigorously true even

⁸ H. Poincaré: "Sur les Principes de la Géométrie, Réponse à M. Russell," *op. cit.*, pp. 85-86, italics in the latter paragraph are mine.

⁹ *Ibid.*, pp. 73-74.

¹ H. Poincaré: *Foundations of Science*, *op. cit.*, p. 82. Cf. also the paper cited in the preceding footnote, where he writes (p. 77): "One would thus have to define distance by measurement" and (p. 78): "The geometric [abstract] distance is thus in need of being defined; and it can be defined only by means of measurement."

though the experimental laws which have determined their adoption are only approximative.²

The only way in which I can construe the latter passage and others like them in the face of our earlier citations from him is by assuming that Poincaré maintained the following: there are practical rather than logical obstacles which frustrate the complete elimination of perturbational distortions, and the resulting vagueness (spread) as well as the finitude of the empirical data provide scope for the exercise of a certain measure of convention in the determination of a metric tensor.

This reading of Poincaré accords with the interpretation of him in L. Rougier's *La Philosophie Géométrique de Henri Poincaré*. Rougier writes:

The conventions fix the language of science which can be indefinitely varied: once these conventions are accepted, the facts expressed by science necessarily are either true or false. . . . Other conventions remain possible, leading to other modes of expressing oneself; but the truth, thus diversely translated, remains the same. One can pass from one system of conventions to another, from one language to another, by means of an appropriate dictionary. The very possibility of a translation shows here the existence of an invariant. . . . Conventions relate to the variable language of science, not to the invariant reality which they express.³

(C) CRITICAL EVALUATION OF EINSTEIN'S CONCEPTION OF
THE INTERDEPENDENCE OF GEOMETRY AND PHYSICS:
PHYSICAL GEOMETRY AS A COUNTER-EXAMPLE TO THE
NON-TRIVIAL D-THESIS.

Einstein has articulated and endorsed Duhem's claim by reference to the special case of testing a hypothesis of physical geometry. In opposition to the Carnap-Reichenbach conception, Einstein maintains⁴ that no hypothesis of physical geometry is *separately* falsifiable, i.e., in isolation from the remainder of

See Append. §9

² *Ibid.*, pp. 64–65. Similar statements are found on pp. 79 and 240.

³ L. Rougier: *La Philosophie Géométrique de Henri Poincaré* (Paris: F. Alcan; 1920), pp. 200–201.

⁴ Cf. A. Einstein: "Reply to Criticisms," *op. cit.*, pp. 676–78.

physics, even though all of the terms in the vocabulary of the geometrical theory, including the term "congruent" for line segments and angles, have been given a specific physical interpretation. And the substance of his argument is briefly the following: In order to follow the practice of ordinary physics and use rigid solid rods as the physical standard of congruence in the determination of the geometry, it is essential to make computational allowances for the thermal, elastic, electromagnetic, and other deformations exhibited by solid rods. The introduction of these corrections is an essential part of the logic of testing a physical geometry.⁵ For the presence of inhomogeneous thermal and other such influences issues in a dependence of the coincidence behavior of transported solid rods on the latter's *chemical composition*, whereas physical geometry is conceived as the system of metric relations exhibited by transported solid bodies independently of their particular chemical composition. The demand for the computational *elimination* of such substance-specific distortions as a prerequisite to the experimental determination of the geometry has a thermodynamic counterpart: the requirement of a means for measuring temperature which does not yield the discordant results produced by expansion thermometers at other than fixed points when different thermometric substances are employed. This thermometric need is fulfilled successfully by Kelvin's thermodynamic scale of temperature.

But Einstein argues that the geometry itself can never be accessible to experimental falsification in isolation from those other laws of physics which enter into the calculation of the corrections compensating for the distortions of the rod. And from this he then concludes that you can always preserve any geometry you like by suitable adjustments in the associated correctional physical laws. Specifically, he states his case in the form of a dialogue in which he attributes his own Duhemian view to Poincaré and offers that view in opposition to Hans Reichenbach's

⁵ For a very detailed treatment of the relevant computations, see B. Weinstein: *Handbuch der Physikalischen Massbestimmungen* (Berlin: Julius Springer), Vol. I (1886), and Vol. II (1888); A. Pérard: *Les Mesures Physiques* (Paris: Presses Universitaires de France; 1955); U. Stille: *Messen und Rechnen in der Physik* (Braunschweig: Vieweg; 1955), and R. Lécercq: *Guide Théorique et Pratique de la Recherche Expérimentale* (Paris: Gauthier-Villars; 1958).

conception, which was discussed in Chapter Three. But I submit that Poincaré's text will *not* bear Einstein's interpretation. For, as we saw in Section B, when speaking of the variations which solids exhibit under distorting influences, Poincaré says "we neglect these variations in laying the foundations of geometry, because, besides their being very slight, they are irregular and consequently seem to us accidental."⁶ I am therefore taking the liberty of replacing the name "Poincaré" in Einstein's dialogue by the term "Duhem and Einstein." With this modification, the dialogue reads as follows:

Duhem and Einstein: The empirically given bodies are not rigid, and consequently can not be used for the embodiment of geometric intervals. Therefore, the theorems of geometry are not verifiable.

Reichenbach: I admit that there are no bodies which can be *immediately* adduced for the "real definition" [i.e., physical definition] of the interval. Nevertheless, this real definition can be achieved by taking the thermal volume-dependence, elasticity, electro- and magneto-striction, etc., into consideration. That this is really and without contradiction possible, classical physics has surely demonstrated.

Duhem and Einstein: In gaining the real definition improved by yourself you have made use of physical laws, the formulation of which presupposes (in this case) Euclidean geometry. The verification, of which you have spoken, refers, therefore, not merely to geometry but to the entire system of physical laws which constitute its foundation. An examination of geometry by itself is consequently not thinkable. —Why should it consequently not be entirely up to me to choose geometry according to my own convenience (i.e., Euclidean) and to fit the remaining (in the usual sense "physical") laws to this choice in such manner that there can arise no contradiction of the whole with experience?⁷

By speaking here of the "real definition" (i.e., the coordinative definition) of "congruent intervals" by the corrected transported rod, Einstein is ignoring that the actual and potential physical meaning of congruence in physics *cannot* be given exhaustively

⁶ H. Poincaré: *The Foundations of Science*, *op. cit.*, p. 76.

⁷ A. Einstein: "Reply to Criticisms," *op. cit.*, pp. 676–78 as modified.

by any *one* physical criterion or test condition. But here as elsewhere in this book, we can safely ignore this open cluster character of the concept of congruence. Our concern as well as Einstein's is merely to single out *one* particular congruence class from among an infinitude of such alternative classes. And as long as our specification of that one chosen class is unambiguous, it is wholly immaterial that there are also other physical criteria (or test conditions) by which it could be specified.

Einstein is making two major points here: First, in obtaining a physical geometry by giving a physical interpretation of the postulates of a formal geometric axiom system, the specification of the physical meaning of such theoretical terms as "congruent," "length," or "distance" is not at all simply a matter of giving an operational definition in the strict sense. Instead, what has been variously called a "rule of correspondence" (Margenau and Carnap), a "coordinative definition" (Reichenbach), an "epistemic correlation" (Northrop), or a "dictionary" (N. R. Campbell) is provided here through the mediation of hypotheses and laws which are collateral to the geometric theory whose physical meaning is being specified. Einstein's point that the physical meaning of congruence is given by the transported rod *as corrected theoretically* for idiosyncratic distortions is an illuminating one and has an abundance of analogues throughout physical theory, thus showing, incidentally, that strictly operational definitions are a rather simplified and limiting species of rules of correspondence. In particular, we see that the physical interpretation of the term "length," which is often adduced as the prototype of all "operational" definitions in Bridgman's sense, is not given operationally in any distinctive sense of that ritually invoked term. Einstein's second claim, which is the cardinal one for our purposes, is that the role of collateral theory in the physical definition of congruence is such as to issue in the following circularity, from which there is no escape, he maintains, short of acknowledging the existence of an a priori element in the sense of the Duhemian ambiguity: the rigid body is not even defined without first decreeing the validity of Euclidean geometry (or of some other particular geometry). For before the corrected rod can be used to make an empirical determination of the *de facto* geometry, the required corrections must be computed via laws, such as those of elasticity, which involve Euclideanly cal-

culated areas and volumes.⁸ But clearly the warrant for thus introducing Euclidean geometry *at this stage* cannot be empirical.

In the same vein, H. Weyl endorses Duhem's position as follows:

Geometry, Mechanics, and Physics form an inseparable theoretical whole. . . .⁹ Philosophers have put forward the thesis that the validity or non-validity of Euclidean geometry cannot be proved by empirical observations. It must in fact be granted that in all such observations essentially physical assumptions, such as the statement that the path of a ray of light is a straight line and other similar statements, play a prominent part. This merely bears out the remark already made above that it is only the whole composed of geometry and physics that may be tested empirically.¹

If Einstein's and Weyl's Duhemian thesis were to prove correct, then it would have to be acknowledged that there is a sense in which physical geometry itself does not provide a geometric characterization of physical reality. For by this characterization we understand the articulation of the system of relations obtaining between bodies and transported solid rods quite apart from their substance-specific distortions. And to the extent to which physical geometry is a priori in the sense of the Duhemian ambiguity, there is an ingression of a priori elements into physical theory to take the place of distinctively geometric gaps in our knowledge of the physical world.

I now wish to set forth my doubts regarding the soundness of Einstein's geometrical form of the D-thesis by demonstrating the separate falsifiability of the geometric hypothesis H. And I shall do so in two parts, the first of which deals with the simplified case in which effectively no deforming influences are present in a certain region whose geometry is to be ascertained. Having argued in Section (A) that the non-trivial D-thesis is a *non-sequitur*, my present aim is to show by geometric counter-example that it is also false.

⁸ Cf. I. S. Sokolnikoff: *Mathematical Theory of Elasticity* (New York: McGraw-Hill Book Company; 1946), and S. Timoshenko and J. N. Goodier: *Theory of Elasticity* (New York: McGraw-Hill Book Company; 1951).

⁹ H. Weyl: *Space-Time-Matter* (New York: Dover Publications, Inc.; 1950), p. 67.

¹ *Ibid.*, p. 93.

If we are confronted with the problem of the falsifiability of the geometry ascribed to a region which is effectively free from deforming influences, then the correctional physical laws play no role as auxiliary assumptions, and the latter reduce to the claim that the region in question is, in fact, effectively free from deforming influences. And if such freedom can be affirmed without presupposing collateral theory, then the geometry alone rather than only a wider theory in which it is ingredient will be falsifiable. By contrast, if collateral theory were presupposed here, then Duhem and Einstein might be able to adduce its modifiability to support their claim that the geometry itself is not separately falsifiable.

Specifically, they might argue then that the collateral theory could be modified such that the region then turns out not to be free from deforming influences with resulting inconclusive falsifiability of the geometry. The question is therefore whether freedom from deforming influences can be asserted and ascertained independently of (sophisticated) collateral theory. My answer to this question is "Yes." For quite independently of the conceptual elaboration of such physical magnitudes as temperature, whose constancy would characterize a region free from deforming influences, the absence of perturbations is certifiable for the region as follows: two solid rods of very different chemical constitution which coincide at one place in the region will also coincide everywhere else in it (independently of their paths of transport). It would not do for the Duhemian to object here that the certification of two solids as quite different chemically is theory-laden to an extent permitting him to uphold his thesis of the inconclusive falsifiability of the geometry. For suppose that observations were so ambiguous as to permit us to assume that two solids which appear strongly to be chemically different are, in fact, chemically identical in all relevant respects. If so rudimentary an observation were thus ambiguous, then no observation could ever possess the required univocity to be incompatible with an observational consequence of a total theoretical system. And if that were the case, Duhem could hardly avoid the following conclusion: "observational findings are always so unrestrictedly ambiguous as not to permit even the refutation of any given total theoretical system." But such a result would

be tantamount to the absurdity that any total theoretical system can be espoused as true a priori.

By the same token, incidentally, I cannot see what methodological safeguards would prevent Quine from having to countenance such an outcome within the framework of his D-thesis. In view of his avowed willingness to "plead hallucination" to deal with observations not conforming to the hypothesis that "there are brick houses on Elm Street," one wonders whether he would be prepared to say that *all* human observers who make disconfirming observations on Elm Street are hallucinating. And, if so, why not discount all observations incompatible with an *arbitrary* total theoretical system as hallucinatory? Thus, it would seem that if Duhem is to maintain, as he does, that a total theoretical *system* is refutable by confrontation with observational results, then he must allow that the coincidence of diverse kinds of rods at different places in the region (independently of their paths of transport) is certifiable observationally. Accordingly, the absence of deforming influences is ascertainable *independently* of any assumptions as to the geometry and of other (sophisticated) collateral theory.

Let us now employ our earlier notation and denote the geometry by "H" and the assertion concerning the freedom from perturbations by "A." Then, once we have laid down the congruence definition and the remaining semantical rules, the physical geometry H becomes *separately* falsifiable as an *explanans* of the posited empirical findings O'. It is true, of course, that A is only more or less highly confirmed by the ubiquitous coincidence of chemically different kinds of solid rods. But the inductive risk thus inherent in affirming A does not arise from the alleged inseparability of H and A, and that risk can be made exceedingly small without any involvement of H. Accordingly, the actual logical situation is characterized not by the Duhemian schema but instead by the schema

$$[(H \cdot A) \rightarrow O] \cdot \sim O \cdot A \rightarrow \sim H.$$

It will be noted that we identified the H of the Duhemian schema with the geometry. But since a geometric theory, at least in its synthetic form, can be axiomatized as a conjunction of logically independent postulates, a particular axiomatization of

H could be decomposed logically into various sets of component subhypotheses. Thus, for example, the hypothesis of Euclidean geometry could be stated, if we wished, as the conjunction of two parts consisting respectively of the Euclidean parallel postulate and the postulates of absolute geometry. And the hypothesis of hyperbolic geometry could be stated in the form of a conjunction of absolute geometry and the hyperbolic parallel postulate.

In view of the logically compounded character of a geometric hypothesis, Professor Grover Maxwell has suggested that the Duhemian thesis may be tenable in this context if we construe it as pertaining not to the falsifiability of a geometry as a whole but to the falsifiability of its component subhypotheses in any given axiomatization. There are two ways in which this proposed interpretation might be understood: first, as an assertion that *any one component subhypothesis* eludes conclusive refutation on the grounds that the empirical findings can falsify the set of axioms only as a whole, or second, in any given axiomatization of a physical geometry there exists *at least one component subhypothesis* which eludes conclusive refutation.

The first version of the proposed interpretation will not bear examination. For suppose that H is the hypothesis of Euclidean geometry and that we consider absolute geometry as one of its subhypotheses and the Euclidean parallel postulate as the other. If now the empirical findings were to show, on the one hand, that the geometry is hyperbolic, then indeed absolute geometry would have eluded refutation; but if, on the other hand, the prevailing geometry were to turn out to be spherical, then the mere replacement of the Euclidean parallel postulate by the spherical one could not possibly save absolute geometry from refutation. For absolute geometry alone is logically incompatible with spherical geometry and hence with the posited empirical findings.

If one were to read Duhem as per the very cautious *second* version of Maxwell's proposed interpretation, then our analysis of the logic of testing the geometry of a *perturbation-free* region could not be adduced as having furnished a counter-example to so mild a form of Duhemism. And the question of the validity of this highly attenuated version is thus left open by our analysis without any detriment to that analysis.

We now turn to the critique of Einstein's Duhemian argument

as applied to the empirical determination of the geometry of a region which is subject to deforming influences.

There can be no question that when deforming influences are present, the laws used to make the corrections for deformations involve areas and volumes in a fundamental way (e.g., in the definitions of the elastic stresses and strains) and that this involvement presupposes a geometry, as is evident from the area and volume formulae of differential geometry, which contain the square root of the determinant of the components g_{ik} of the metric tensor.² Thus, the empirical determination of the geometry involves the joint assumption of a geometry and of certain collateral hypotheses. But we see already that this assumption cannot be adequately represented by the conjunction $H \cdot A$ of the Duhemian schema, where H represents the geometry.

Now suppose that we begin with a set of Euclideanly formulated physical laws P_0 in correcting for the distortions induced by perturbations and then use the thus Euclideanly corrected congruence standard for empirically exploring the geometry of space by determining the metric tensor. *The initial stipulational affirmation of the Euclidean geometry G_0 in the physical laws P_0 used to compute the corrections in no way assures that the geometry obtained by the corrected rods will be Euclidean!* If it is non-Euclidean, then the question is: What will be required by Einstein's fitting of the physical laws to preserve Euclideanism and avoid a contradiction of the theoretical system with experience? Will the adjustments in P_0 necessitated by the retention of Euclidean geometry entail merely a change in the dependence of the length assigned to the transported rod on such nonpositional parameters as temperature, pressure, and magnetic field? Or could the putative empirical findings compel that the length of the transported rod be likewise made a nonconstant function of its position and orientation as independent variables in order to square the coincidence findings with the requirement of Euclideanism? The possibility of obtaining non-Euclidean results by measurements carried out in a spatial region uniformly characterized by standard conditions of temperature, pressure, electric and magnetic field strength, etc., shows it to be extremely

² L. P. Eisenhart: *Riemannian Geometry* (Princeton: Princeton University Press; 1949), p. 177.

doubtful, as we shall now show, that the preservation of Euclideanism could always be accomplished short of introducing *the dependence of the rod's length on the independent variables of position or orientation*.

But the introduction of the latter dependence is none other than so radical a change in the meaning of the word "congruent" that this term now denotes a class of intervals different from the original congruence class denoted by it. And such tampering with the semantical anchorage of the word "congruent" violates the requirement of semantical stability, which is a necessary condition for the *non-triviality* of the D-thesis, as we saw in Section A.

Suppose that, relatively to the customary congruence standard, the geometry prevailing in a given region when free from perturbational influences is that of a strongly non-Euclidean space of spatially and temporally constant curvature. Then what would be the character of the alterations in the customary correctional laws which Einstein's thesis would require to assure the Euclideanism of that region relatively to the customary congruence standard under *perturbational* conditions? The required alterations would be *independently falsifiable*, as will now be demonstrated, because they would involve affirming that such coefficients as those of linear thermal expansion *depend on the independent variables of spatial position*. That such a space dependence of the correctional coefficients might well be necessitated by the exigencies of Einstein's Duhemian thesis can be seen as follows by reference to the law of linear thermal expansion. In the usual version of physical theory, the first approximation of that law³ is given in terms of the deviation ΔT from the standard temperature by

$$L = L_0 (1 + \alpha \cdot \Delta T).$$

³ This law is only the first approximation, because the rate of thermal expansion varies with the temperature. The general equation giving the magnitude m_t (length or volume) at a temperature t , where m_0 is the magnitude at 0°C , is

$$m_t = m_0 (1 + \alpha t + \beta t^2 + \gamma t^3 + \dots),$$

where α , β , γ , etc., are empirically determined coefficients [Cf. *Handbook of Chemistry and Physics* (Cleveland: Chemical Rubber Publishing Company; 1941), p. 2194]. The argument which is about to be given by reference to the approximate form of the law can be readily generalized to forms of the law involving more than one coefficient of expansion.

If Einstein is to guarantee the Euclideanism of the region under discussion by means of logical devices that are consonant with his thesis, and if our region is subject only to thermal perturbations for some time, then we are confronted with the following situation: unlike the customary law of linear thermal expansion, the revised form of that law needed by Einstein will have to bear the twin burden of effecting both of the following two kinds of superposed corrections: first, the *changes* in the lengths ascribed to the transported rod in different positions or orientations which would be required even if our region were everywhere at the standard temperature, merely for the sake of rendering Euclidean its otherwise non-Euclidean geometry, and second, corrections compensating for the effects of the de facto deviations from the standard temperature, these corrections being the sole onus of the usual version of the law of linear thermal expansion. What will be the consequences of requiring the revised version of the law of thermal elongation to implement the *first* of these two kinds of corrections in a context in which the deviation ΔT from the standard temperature is the *same* at some *different* points of the region (e.g., $\Delta T = 0$), that temperature deviation having been measured in the manner chosen by the Duhemian? Specifically, what will be the character of the coefficients α of the revised law of thermal elongation under the posited circumstances, if Einstein's thesis is to be implemented by effecting the *first* set of corrections? Since the new version of the law of thermal expansion will then have to guarantee that the lengths L assigned to the rod at the various points of *equal* temperature T *differ* appropriately, it would seem clear that logically possible empirical findings could compel Einstein to make the coefficients α of solids *depend* on the *space coordinates*.

But such a spatial dependence is *independently falsifiable*: comparison of the thermal elongations of an aluminum rod, for example, with an invar rod of essentially zero α by, say, the Fizeau method might well show that the α of the aluminum rod is a characteristic of aluminum which is not dependent on the space coordinates. And even if it were the case that the α 's are found to be space dependent, how could Duhem and Einstein assure that this space dependence would have the particular functional form required for the success of their thesis?

We see that the required resort to the introduction of a spatial

dependence of the thermal coefficients might well not be open to Einstein. Hence, in order to retain Euclideanism, it would then be necessary to remetrize the space in the sense of abandoning the customary definition of congruence, entirely apart from any consideration of idiosyncratic distortions and even after correcting for these in some way or other. But this kind of remetrization, though entirely admissible in other contexts, does not provide the requisite support for Einstein's Duhemian thesis! For Einstein offered it as a criticism of Reichenbach's conception. And hence it is the avowed onus of that thesis to show that the geometry by itself cannot be held to be empirical, i.e., separately falsifiable, even when, with Reichenbach, we have sought to assure its empirical character by choosing and then adhering to the usual (standard) definition of spatial congruence, which excludes resorting to such remetrization.

Thus, there may well obtain observational findings (O' , expressed in terms of a particular definition of congruence (e.g., the customary one), which are such that there does not exist any non-trivial set A' of auxiliary assumptions capable of preserving the Euclidean H in the face of O' . And this result alone suffices to invalidate the Einsteinian version of Duhem's thesis to the effect that any geometry, such as Euclid's, can be preserved in the face of any experimental findings which are expressed in terms of the customary definition of congruence.

It might appear that my geometric counterexample to the Duhemian thesis of unavoidably contextual falsifiability of an *explanans* is vulnerable to the following criticism: "To be sure, Einstein's geometric articulation of that thesis does not leave room for saving it by resorting to a remetrization in the sense of making the length of the rod vary with position or orientation even *after* it has been corrected for idiosyncratic distortions. But why saddle the Duhemian thesis as such with a restriction peculiar to Einstein's particular version of it? And thus why not allow Duhem to save his thesis by countenancing those alterations in the congruence definition which are remetrizations?"

My reply is that to deny the Duhemian the invocation of such an alteration of the congruence definition in this context is not a matter of gratuitously requiring him to justify his thesis within the confines of Einstein's particular version of that thesis; instead,

the imposition of this restriction is entirely legitimate here, and the Duhemian could hardly wish to reject it as unwarranted. For it is of the essence of Duhem's contention that H (in this case, Euclidean geometry) can always be preserved not by tampering with the principal semantical rules (interpretive sentences) linking H to the observational base (i.e., the rules specifying a particular congruence class of intervals, etc.), but rather by availing oneself of the alleged *inductive latitude* afforded by the ambiguity of the experimental evidence to do the following: (a) leave the factual commitments of H essentially *unaltered* by retaining both the statement of H and the principal semantical rules linking its terms to the observational base, and (b) replace the set A by A' such that A and A' are logically incompatible under the hypothesis H. *The qualifying words "principal" and "essential" are needed here in order to obviate the possible objection that it may not be logically possible to supplant the auxiliary assumptions A by A' without also changing the factual content of H in some respect.* Suppose, for example, that one were to abandon the optical hypothesis A that light will require equal times to traverse congruent closed paths in an inertial system in favor of some rival hypothesis. Then the semantical linkage of the term "congruent space intervals" to the observational base is changed to the extent that this term no longer denotes intervals traversed by light in equal round-trip times. But such a change in the semantics of the word "congruent" is innocuous in this context, since *it leaves wholly intact the membership of the class of spatial intervals that is referred to as a "congruence class."* In this sense, then, the modification of the optical hypothesis leaves intact both the "principal" semantical rule governing the term "congruent" and the "essential" factual content of the geometric hypothesis H, which is predicated on a particular congruence class of intervals. That "essential" factual content is the following: relatively to the congruence specified by unperturbed transported rods—among other things—the geometry is Euclidean.

Now, the essential factual content of a geometrical hypothesis can be changed either by preserving the original statement of the hypothesis while changing one or more of the principal semantical rules or by keeping all of the semantical rules intact and suitably changing the statement of the hypothesis. We can see,

therefore, that the retention of a Euclidean H by the device of changing through remetrization the semantical rule governing the meaning of "congruent" (for line segments) effects a retention not of the essential *factual commitments* of the original Euclidean H but only of its *linguistic trappings*. That the thus "preserved" Euclidean H actually repudiates the essential factual commitments of the *original* one is clear from the following: the *original* Euclidean H had asserted that the coincidence behavior common to all kinds of solid rods is Euclidean, *if* such transported rods are taken as the physical realization of congruent intervals; but the Euclidean H which survived the confrontation with the posited empirical findings only by dint of a *remetrization* is predicated on a denial of the very assertion that was made by the original Euclidean H, which it was to "preserve." It is as if a physician were to endeavor to "preserve" an a priori diagnosis that a patient has acute appendicitis in the face of a negative finding (yielded by an exploratory operation) as follows: he would redefine "acute appendicitis" to denote the healthy state of the appendix!

Hence, the confines within which the Duhemian must make good his claim of the preservability of a Euclidean H do not admit of the kind of change in the congruence definition which alone would render his claim tenable under the assumed empirical conditions. Accordingly, the geometrical critique of Duhem's thesis given here does not depend for its validity on restrictions peculiar to Einstein's version of it.

Even apart from the fact that Duhem's thesis precludes resorting to an alternative metrization to save it from refutation in our geometrical context, the very feasibility of alternative metrizations is vouchsafed not by any general Duhemian considerations pertaining to the logic of falsifiability but by a property peculiar to the subject matter of geometry (and chronometry): the latitude for convention in the ascription of the spatial (or temporal) equality relation to intervals in the continuous manifolds of physical space (or time).

But what of the possibility of actually extricating the unique underlying geometry (to within experimental accuracy) from the network of hypotheses which enter into the testing procedure?

That contrary to Duhem and Einstein, the geometry itself may well be empirical, once we have renounced the kinds of alterna-

tive congruence definitions employed by Poincaré is seen from the following possibilities of its successful empirical determination. After assumedly obtaining a non-Euclidean geometry G_1 from measurements with a rod corrected on the basis of Euclideanly formulated physical laws P_0 , we can revise P_0 so as to conform to the non-Euclidean geometry G_1 just obtained by measurement. This retroactive revision of P_0 would be effected by recalculating such quantities as areas and volumes on the basis of G_1 and changing the functional dependencies relating them to temperatures and other physical parameters. Let us denote by " P'_1 " the set of physical laws P resulting from thus revising P_0 to incorporate the geometry G_1 . Since various physical magnitudes ingredient in P'_1 involve lengths and durations, we now use the set P'_1 to correct the rods (and clocks) with a view to seeing whether rods and clocks thus corrected will reconfirm the set P'_1 . If not, modifications need to be made in this set of laws so that the functional dependencies between the magnitudes ingredient in them reflect the new standards of spatial and temporal congruence defined by P'_1 -corrected rods and clocks. We may thus obtain a new set of physical laws P_1 . Now we employ this set P_1 of laws to correct the rods for perturbational influences and then determine the geometry with the thus corrected rods. Suppose the result is a geometry G_2 different from G_1 . Then *if*, upon repeating this two step process several more times, there is convergence to a geometry of constant curvature we must continue to repeat the two step process an additional finite number of times until we find the following: the geometry G_n ingredient in the laws P_n providing the basis for perturbation-corrections is indeed the same (to within experimental accuracy) as the geometry obtained by measurements with rods that have been corrected via the set P_n .

If there is such convergence at all, it will be to the same geometry G_n even if the physical laws used in making the *initial* corrections are not the set P_0 , which presupposes Euclidean geometry, but a different set P based on some non-Euclidean geometry or other. That there can exist only one such geometry of constant curvature G_n would seem to be guaranteed by the identity of G_n with the unique underlying geometry G_1 characterized by the following properties: (1) G_1 would be exhibited by the coincidence behavior of a transported rod if the *whole* of the

space were actually free of deforming influences, (2) G_i would be obtained by measurements with rods corrected for distortions on the basis of physical laws P_i presupposing G_i , and (3) G_i would be found to prevail in a given relatively small, perturbation-free region of the space quite independently of the assumed geometry ingredient in the correctional physical laws. Hence, *if* our method of successive approximation does converge to a geometry G_n of *constant* curvature, then G_n would be this unique underlying geometry G_i . And, in that event, we can claim to have found empirically that G_i is indeed the geometry prevailing in the entire space which we have explored.

But what if there is no convergence? It might happen that whereas convergence would obtain by starting out with corrections based on the set P_o of physical laws, it would *not* obtain by beginning instead with corrections presupposing some particular non-Euclidean set P or vice versa: just as in the case of Newton's method of successive approximation,⁴ there are conditions, as A. Suna has pointed out to me, under which there would be no convergence. We might then nonetheless succeed as follows in finding the geometry G_i empirically, *if* our space is one of constant curvature.

See Append. §8

The geometry G_r resulting from measurements by means of a corrected rod is a single-valued function of the geometry G_a assumed in the correctional physical laws, and a Laplacian demon having sufficient knowledge of the facts of the world would know this function $G_r = f(G_a)$. Accordingly, we can formulate the problem of determining the geometry empirically as the problem of finding the point of intersection between the curve representing this function and the straight line $G_r = G_a$. That there exists one and only one such point of intersection follows from the existence of the geometry G_i defined above, provided that our space is one of constant curvature. Thus, what is now needed is to make determinations of the G_r corresponding to a number of geometrically different sets of correctional physical laws P_n , to draw the most reasonable curve $G_r = f(G_a)$ through this finite number of points (G_a, G_r) , and then to find the point of intersection of this curve and the straight line $G_r = G_a$.

Whether this point of intersection turns out to be the one repre-

⁴ R. Courant: *Vorlesungen über Differential und Integralrechnung* (Berlin: Julius Springer; 1927), Vol. I, p. 286.

senting Euclidean geometry or not is beyond the reach of our conventions, barring a remetrization. And thus the least that we can conclude is that since empirical findings can greatly narrow down the range of uncertainty as to the prevailing geometry, there is no assurance of the latitude for the choice of a geometry which Einstein takes for granted in the manner of the D-thesis. Einstein's Duhemian position would appear to be safe from this further criticism only if our proposed method of determining the geometry by itself empirically cannot be generalized in some way to cover the general relativity case of a space of variable curvature and if the latter kind of theory turns out to be true. The extension of our method to the case of a geometry of variable curvature is not simple. For in that case, the geometry G is no longer represented by a single scalar given by a Gaussian curvature, and our graphical method breaks down.

If my proposed method of escaping from the web of the Duhemian ambiguity were shown to be unsuccessful, and if there should happen to be no other scientifically viable means of escape, then, it seems to me, we would unflinchingly have to resign ourselves to this relatively unmitigable type of uncertainty. No, says the philosopher Jacques Maritain, who enticingly beckons us to take heart. The *scientific* elusiveness of the correct geometric description of external reality must *not* lead us to suppose, he tells us, that philosophy, when divorced from mathematical physics, cannot rescue us from the labyrinth of the Duhemian perplexity and unveil for us the structure of what he calls *ens geometricum reale* (real geometrical being).⁵ As against Maritain's conception of the capabilities of philosophy as an avenue of cognition, I wish to uphold the following excellent declaration by Professor Bridgman: "The physicist emphatically would *not* say that his knowledge presumptively will not lead to a full understanding of reality for the reason that there are other kinds of knowledge than the knowledge in which he deals."⁶ To justify my endorsement of Professor Bridgman's statement in this

⁵ J. Maritain: *The Degrees of Knowledge* (London: G. Bles Company; 1937), p. 207. A new translation of this work was published in New York in 1959 by Charles Scribner's Sons; all page references here are to the 1937 translation.

⁶ P. W. Bridgman: "The Nature of Physical 'Knowledge,'" in: L. W. Friedrich (ed.) *The Nature of Physical Knowledge* (Bloomington: Indiana University Press; 1960), p. 21.

context, I shall give a brief critique of Maritain's philosophy of geometry as presented in his book *The Degrees of Knowledge*.⁷

I have selected Maritain's views for rebuttal because they typify the conception of those who believe that the philosopher as such has at his disposal means for fathoming the structure of external reality which are not available to the scientist. In outline, Maritain endeavors to justify this idea in regard to geometry along the following lines. Says he: "There is no clearer word than the word reality, which means *that which is*. . . . What is meant when it is asked whether real space is euclidean or non-euclidean . . . ?"⁸ To prepare for his answer to this question, he explains the following: "The word *real* has not the same meaning for the philosopher, the mathematician and the physicist." . . . For the physicist a space is 'real' when the geometry to which it corresponds permits of the construction of a physico-mathematical universe which coherently and completely symbolizes physical phenomena, and where all our graduated readings find themselves 'explained.' And it is obvious that from this point of view no space of any kind holds any sort of privileged position.¹ But . . . the question is to know what is real space in the philosophical meaning of the word, i.e., as a 'real' entity . . . designating an object of thought capable of an extra-mental existence. . . ."² One is immediately puzzled as to how Maritain conceives that his distinction between physically real space and the philosophically real space which is avowedly extra-mental is not an empty distinction without a difference. And, instead of being resolved, this puzzlement only deepens when he tells us that by the *extra-mental geometric* features of existing bodies he understands "those properties which the mind recognizes in the elimination of all the physical."³ But let us suspend judgment concerning this difficulty and see whether it is not cleared up by his treatment of the following question posed by him: How are we to know whether it is Euclidean geometry or one of the non-Euclidean geometries that represents the structure of the philosophically real, i.e.,

⁷ J. Maritain: *The Degrees of Knowledge*, op. cit., pp. 201-12.

⁸ *Ibid.*, p. 201.

⁹ *Ibid.*

¹ *Ibid.*, p. 202.

² *Ibid.*, p. 203.

³ *Ibid.*, p. 207.

extra-mental or external space?⁴ In regard to this question, he makes the following assertions:

First, the capabilities of physical measurements to yield the answer to the question are nil,⁵ because a geometry is presupposed in the theory of our measuring instruments which forms the basis of corrections for "accessory variations due to various physical circumstances."⁶ Of course, we recognize this contention to be a strong form of the Duhemian one, although Maritain does not refer to Duhem.

Second, the several non-Euclidean geometries depend for their consistency on their formal translatability into Euclidean geometry. This translation is effected by providing a Euclidean model of the particular non-Euclidean geometry in the sense of embedding an appropriately curved non-Euclidean surface in the three-dimensional Euclidean space. And the privileged position which Euclidean geometry enjoys as the underwriter of the consistency of the non-Euclidean geometries thus issues in a correlative dependence of the intuitability of the non-Euclidean geometries on the primary intuitability of the Euclideanism of the embedding three-dimensional hyperspace.⁷

Using the twin arguments from consistency and intuitability, Maritain then reaches the following final conclusion:

The non-Euclidean spaces can then without the least intrinsic contradiction be the object of consideration by the mind, but there would be a contradiction in supposing their existence outside the mind, and thereby suppressing, for their benefit, the existence of the foundation on which the notion of them is based.

Either way we are thus led to admit, despite the use which astronomy makes of them, that these non-euclidean spaces are rational [i.e., *purely mental*] beings; and that the *geometric* properties of existing bodies, those properties which the mind recognizes in the elimination of all the physical, are those which characterize euclidean space. For philosophy it is euclidean space which appears as an *ens geometricum reale*.⁸

⁴ Cf. *ibid.*, p. 204.

⁵ Cf. *ibid.*, p. 204.

⁶ *Ibid.*, p. 205.

⁷ Cf. *ibid.*, pp. 202, 205–206.

⁸ *Ibid.*, pp. 206–7.

I submit that Maritain's thesis is unsound in its entirety and can be completely refuted as follows:

First, as Hilbert and Bernays have explained,⁹ the consistency of the Euclidean axiom system is not vouchsafed by its intuitive plausibility as an adequate description of the space of our immediate physical environment. Instead, we establish the consistency of Euclidean geometry by providing a model of the formal Euclidean postulates in the domain of real numbers in the manner of analytic geometry.¹ Now, Maritain overlooks that precisely the same procedure of providing a real number model can be used to establish the internal consistency of the various non-Euclidean geometries without the mediation of a prior translation into Euclidean geometry (except possibly in an irrelevant heuristic sense). And he is misled by the fact that, historically, the consistency of the several non-Euclidean geometries was established by means of a translation into Euclidean geometry, as for example in Klein's relative consistency proof of hyperbolic geometry via a model furnished by the interior of a circle in the Euclidean plane. For surely the temporal priority of Euclidean geometry inherent in the historical circumstances of our discovery of the internal consistency of the various non-Euclidean geometries hardly serves to establish the logical primacy of Euclidean geometry as the sole guarantee of their consistency. And Maritain's error on this count is only compounded by his intuitability argument for the uniqueness of Euclideanism as the only possible structure of extra-mental reality. The latter argument is vitiated by the inveterate error of being victimized by the misleading connotation of *embedding in a Euclidean hyperspace*, which is possessed by the terms "curved space" and "curvature of a surface." This connotation springs from unawareness that the Gaussian curvature of a 2-space and the Riemannian curvatures for the various orientations at points of a 3-space are intrinsically definable and discernible properties of these spaces, requiring no embedding. Moreover, Maritain overlooks here that even when the consistency proof of hyperbolic geometry, for example, is

⁹ D. Hilbert and P. Bernays: *Grundlagen der Mathematik* (Berlin: Julius Springer; 1934), Vol. I, §1, section A, pp. 2-3.

¹ Cf. L. P. Eisenhart: *Coordinate Geometry* (New York: Dover Publications, Inc.; 1960), Appendix to Chap. 1, pp. 279-92.

given on the basis of Euclidean geometry—which we saw is quite unnecessary—this can be accomplished without embedding, as in the case of the aforementioned *two*-dimensional Klein model, just as readily as by Beltrami's procedure of embedding a surface of constant negative Gaussian curvature (containing singular lines) in Euclidean 3-space.

Lastly, it can surely not be maintained that "the *geometric* properties of existing bodies" are "those properties which the mind recognizes in the elimination of all the physical." For, in that case, geometry would be the study of purely imagined thought-objects, which will, of course, turn out to have Euclidean properties, if Maritain's imagination thus endows them. And the geometry of such an imagined space could then not qualify as the geometry of Maritain's real or extra-mental space. The geometric theory of external reality does indeed abstract from a large class of physical properties in the sense of being the metrical study of the coincidence behavior of transported solids independently of the solids' substance-specific physical properties. But this kind of abstracting does not deprive metrical coincidence behavior of its physicality. And if the methods of the physicist cannot fathom the laws of that behavior, then certainly no other kind of intellectual endeavor will succeed in doing so.

It is true, of course, that even apart from experimental errors, not to speak of quantum limitations on the accuracy with which the metric tensor of *space-time* can be meaningfully ascertained by measurement,² no *finite* number of data can uniquely determine the functions constituting the representations g_{ik} of the metric tensor in any given coordinate system. But the criterion of *inductive* simplicity which governs the free creativity of the geometer's imagination in his choice of a particular metric tensor here is the same as the one employed in theory formation in any of the non-geometrical portions of empirical science. And choices made on the basis of such inductive simplicity are in principle true or false, unlike those springing from considerations of descriptive simplicity, which merely reflect conventions.

² E. P. Wigner: "Relativistic Invariance and Quantum Phenomena," *Reviews of Modern Physics*, Vol. XXIX (1957), pp. 255–68, and H. Salecker and E. P. Wigner: "Quantum Limitations of the Measurement of Space-Time Distances," *The Physical Review*, Vol. CIX (1958), pp. 571–77.

Chapter 5

EMPIRICISM AND THE GEOMETRY OF VISUAL SPACE

A very brief review of the account of our knowledge of visual space given by Carnap, Helmholtz, and Reichenbach will precede the discussion of some problems posed by very recent *experimental* studies of the geometry of *visual* space.

Distinguishing the space of physical objects from the space of visual experience ("Anschauungsraum"), Carnap sided with empiricism even in his earliest work to the extent of maintaining that the topology of *physical* space is known a posteriori and that the coincidence relations among points disclosed by experience yield a unique metrization for that space once a specific coordinative definition of congruence has been chosen freely.¹ But the neo-Kantian *parti pris* of that period enters in his epistemological interpretation of the axioms governing the topology of *visual* space: "Experience does not provide the justification for them, the axioms are . . . independent of the 'quantity of experience,' i.e., knowledge of them does not, as in the case of a posteriori propositions, become ever more reliable through multiply repeated experience. For, as Husserl has shown, we are dealing here not with facts in the sense of empirically ascertained realities but rather with the essence ("eidos") of certain presentations whose special nature can be grasped in a single immediate ex-

¹ R. Carnap: *Der Raum*, *op. cit.*, pp. 39, 45, 54, 63.

perience."² Reminding us of Kant's distinction between knowledge acquired "with" experience, on the one hand, and "from" experience, on the other, the early Carnap classified these axioms as synthetic a priori propositions in that philosopher's sense.

This theory of the phenomenological a priori was a stronger version of Helmholtz's claim that "space can be transcendental [a priori] while its axioms are not."³ For Helmholtz's concession to Kantianism was merely to regard an *amorphous* visual extendedness as an a priori condition of spatial experience⁴ while proclaiming the a posteriori character of the topological and metrical articulations of that extendedness on the basis of his pioneering method of imagining ("sich ausmalen") the specific sensory contents we would have in worlds having alternative spatial structures.⁵

The phenomenological a priori will not do, however, as an account of our knowledge of the properties of visual space. For it is an *empirical* fact that the experiences resulting from ocular activity have the indefinable attribute which is characteristic of visual extendedness rather than that belonging to tactile explorations or to those experiences that would issue from our possession

² *Ibid.*, p. 22. Cf. also p. 62. For a more recent defense of the thesis that "there are synthetic a priori judgments of spatial intuition," cf. K. Reide-meister: "Zur Logik der Lehre vom Raum," *Dialectica*, Vol. VI (1952), p. 342. For a discussion of related questions, see P. Bernays: "Die Grundbegriffe der reinen Geometrie in ihrem Verhältnis zur Anschauung," *Naturwissenschaften*, Vol. XVI (1928).

³ H. von Helmholtz: *Schriften zur Erkenntnistheorie*, P. Hertz and M. Schlick (eds.), (Berlin: Julius Springer; 1921), p. 140.

⁴ *Ibid.*, pp. 2, 70, 121-22, 140-42, 144-45, 147-48, 152, 158, 161-62, 163, 168, 172, 174. Helmholtz attempts to characterize the distinctive attribute of space, not possessed by other tri-dimensional manifolds, in the following way: "in space, the distance between two points on a vertical can be compared to the horizontal distance between two points on the floor, because a measuring device can be applied successively to these two pairs of points. But we cannot compare the distance between two tones of equal pitch and differing intensity with that between two tones of equal intensity and differing pitch" (*ibid.*, p. 12). Schlick, however, properly notes in his commentary (*ibid.*, p. 28) that this attribute is necessary but not sufficient to render the distinctive character of space.

⁵ *Ibid.*, pp. 5, 22, 164-65. Cf. also K. Gerhardt's papers: "Nichteuklidische Kinematographie," *Naturwissenschaften*, Vol. XX (1932), p. 925, and "Nichteuklidische Anschauung und optische Täuschungen," *Naturwissenschaften*, Vol. XXIV (1936), p. 437.

of a sense organ responding to magnetic disturbances. In the class of all logically possible experiences, the "Wesensschau" provided by our ocular activity must be held to give rise to empirical knowledge. For the only way to assure a priori that all future deliverances of our eyes will possess the characteristic attribute which Husserl would have us ascertain in a single glance is by resorting to a covert tautology via refusing to call the resulting knowledge "knowledge of visual space," unless it possesses that attribute.

Reichenbach made a particularly telling contribution to the disintegration of the Kantian *metrical* a priori of *visual* space by showing that such intuitive compulsion as inheres in the Euclideanism of that space derives from facts of logic in which the Kantian interpretation cannot find a last refuge and that the counter-intuitiveness of non-Euclidean relations is merely the result of both ontogenetic and phylogenetic adaptation to the Euclidicity of the physical space of ordinary life.⁶

In very recent years, experimental mathematico-optical researches by R. K. Luneburg⁷ and A. A. Blank⁸ have even led these authors to contend that although the *physical* space in which sensory depth perception by binocular vision is effective is Euclidean, the binocular *visual* space resulting from psychometric coordination possesses a Lobatchevskian *hyperbolic* geometry of constant curvature. This contention suggests several questions.

⁶ H. Reichenbach: *The Philosophy of Space and Time*, *op. cit.*, pp. 32-34 and 37-43.

⁷ R. K. Luneburg: *Mathematical Analysis of Binocular Vision* (Princeton: Princeton University Press; 1947), and "Metric Methods in Binocular Visual Perception," in *Studies and Essays*, Courant Anniversary Volume (New York: Interscience Publishers, Inc.; 1948), pp. 215-39.

⁸ A. A. Blank: "The Luneburg Theory of Binocular Visual Space," *Journal of the Optical Society of America*, Vol. XLIII (1953), p. 717; "The non-Euclidean Geometry of Binocular Visual Space," *Bulletin of the American Mathematical Society*, Vol. LX (1954), p. 376; "The Geometry of Vision," *The British Journal of Physiological Optics*, Vol. XIV (1957), p. 154; "The Luneburg Theory of Binocular Perception," in S. Koch (ed.) *Psychology, A Study of a Science* (New York: McGraw-Hill Book Company, Inc.; 1958), Study I, Vol. I, Part III, Sec. A. 2; "Axiomatics of Binocular Vision. The Foundations of Metric Geometry in Relation to Space Perception," *Journal of the Optical Society of America*, Vol. XLVIII (1958), p. 328, and "Analysis of Experiments in Binocular Space Perception," *Journal of the Optical Society of America*, Vol. XLVIII (1958), p. 911.

The first of these is how human beings manage to get about so easily in a Euclidean physical environment even though the geometry of visual space is presumably hyperbolic. Blank suggests the following as a possible answer to this question: First, man's motor adjustment to his physical environment does not draw on *visual* data alone; moreover, these do contribute physically true information, since they supply a good approximation to the relative directions of objects and since the mapping of physical onto visual space preserves the topology (though not the metric) of physical space, thereby enabling man to control his motor responses by feedback, as in the parking of a car or threading the eye of a needle; and second, the thesis of the hyperbolicity of visual space rests on data obtained under experimental conditions which are far more restrictive than those accompanying ordinary visual experience. Under ordinary conditions, we secure depth perception by relying on the coordination of our two ocular images which we have learned in the past in the usual contexts. But in order to ascertain the laws of merely *one* of the sources of spatial information—stereoscopic depth perception alone—the experimenters of the Luneburg-Blank theory endeavored to deny their subjects precisely that contextual reliance: there were no guideposts of perspectives and familiar objects whose positions the subject had determined by tactile means, the only visible objects being isolated point lights in an otherwise completely dark room; in fact, the subject was not even allowed to move his head to make judgments by parallax. Since these contextual guideposts are also available in monocular vision, the experimenters assumed that they play no part in the innate physiological processes governing the distinctive sensations of three-dimensional space which are obtained binocularly.

Several additional questions arise in regard to the Luneburg theory upon going beyond its own restricted objectives of furnishing an account of binocular visual perception and attempting to incorporate its thesis of the non-Euclidean structure of visual space in a comprehensive theory of spatial *learning*: (1) how is man able to arrive at a rather correct apprehension of the Euclidean metric relations of his environment by the use of a physiological instrument whose deliverances are claimed to be non-

Euclidean? (2) how can students be taught *Euclidean* geometry by *visual* methods, methods which certainly convey more than the topology of Euclidean space and whose success is therefore not explained by the fact that the purportedly hyperbolic visual space preserves the topology of Euclidean physical space? (3) if we have literally been seeing one of the non-Euclidean geometries of constant negative Gaussian curvature all along, why did it require two thousand years of research in axiomatics even to *conceive* these geometries, the Euclideanism of physical space being affirmed throughout this period? (4) why did such thinkers as Helmholtz and Poincaré first have to retrain their *Anschauung* conceptually in a counterintuitive direction before achieving a ready pictorialization of the Lobatchevski-Bolyai world, a feat which very few can duplicate even now? (5) if we took two groups of school children of equal intelligence and without prior formal geometrical education and taught Euclid to one group while teaching Lobatchevski-Bolyai to the other, why is it the case (if indeed that is the case!) that, in all probability, the first group would exhibit a far better mastery of their material?

The need to answer these questions becomes even greater, if we assume that our ideas concerning the geometry of our immediate physical environment are formed, in the first instance, not by the physical geometry of yardsticks or by the formal study of Euclidean geometry but rather by the psychometry of our visual sense data.

A. A. Blank, to whom the writer submitted these questions, has suggested that these questions may have answers which lie in part along the following lines: First, man has to learn the significance of ever-changing patterns of visual sensations for the metric of *physical* space by *discounting* much of the psychometry of visual sensation, thereby developing the habit of not being very perceptive of the metrical details of his visual experience. Thus, we learn before adulthood to associate with the non-rigid sequence of visual sensations corresponding to viewing a chair in various positions and contexts the attribute of physical rigidity, generally ignoring all but those aspects of the changing appearances that can serve as a basis for action. In fact, laboratory findings show that for any physical configuration whatever, there are

an infinity of others which give the same binocular clues.⁹ Since we retain those aspects of visual experience which enable us to place objects in the contexts useful for action, Euclidean relations can be more readily pictured (though not actually seen or made visible) than those of Lobatchevski; second, those geometrical judgments disclosed by binocular perception which are *common* to both Euclidean and hyperbolic geometry¹ will be true physically as well.

Moreover, there are certain small two-dimensional elements of visual space which are essentially isometric with the corresponding elements of the Euclidean space of physical stimuli. For example, in a plane parallel to the line joining the rotation centers of the eyes, physical metric relations are seen undistorted in the vicinity of a point at the base of the perpendicular to the plane from a point located half way between the eyes. We can therefore obtain first-order visual approximations to the physical Euclidean geometry from viewing small diagrams frontally in this way. In a like manner, we can understand how the concept of similar figures, which is uniquely characteristic of Euclidean geometry among spaces of *constant* curvature, can be conveyed in the context of a non-Euclidean visual geometry: all Riemannian geometries are locally Euclidean, thus possessing a group of similarity transformations in the small; third, the presumed greater ease with which students would master Euclid than Lobatchevski is due to the greater analytical simplicity of the numerical relations of Euclidean geometry.

⁹ Cf. A. A. Blank: "The Luneburg Theory of Binocular Visual Space," *op. cit.*, pp. 721-22, and L. H. Hardy, G. Rand, M. C. Rittler, A. A. Blank and P. Boeder: *The Geometry of Binocular Space Perception* (New York: Columbia University College of Physicians and Surgeons; 1953), pp. 15ff. and 39ff.

¹ For the axioms of the so-called "absolute" geometry relevant here, see R. Baldus: *Nichteuklidische Geometrie*, edited by F. Löbell (3rd revised edition; Berlin: Walter de Gruyter and Company; 1953), *Sammlung Götschen*, Vol. CMLXX, Chap. ii.

Chapter 6

A detailed clarification and correction of this chapter is given in §10 of the Appendix and on pp. 532–533 of ch. 16

THE RESOLUTION OF ZENO'S METRICAL PARADOX OF EXTENSION FOR THE MATHEMATICAL CONTINUA OF SPACE AND TIME

It is a commonplace in the analytic geometry of physical space and time that an *extended* straight line segment, having positive length, is treated as “consisting of” *unextended* points, each of which has zero length. Analogously, time intervals of positive duration are resolved into instants, each of which has zero duration.

Ever since some of the Greeks defined a point as “that which has no part,”¹ philosophers and mathematicians have questioned the consistency of conceiving of an extended continuum as an aggregate of unextended elements. On the long list of investigators who have examined this question in the context of the specific mathematical and philosophical theories of their time, we find such names as Zeno of Elea,² Aristotle,³ Cavalieri,⁴ Tacquet,⁵

¹ This definition is given in Euclid: *The Thirteen Books of Euclid's Elements*, trans. T. L. Heath (Cambridge: Cambridge University Press; 1926), p. 153.

² S. Luria: “Die Infinitesimaltheorie der antiken Atomisten,” *Quellen und Studien zur Geschichte der Mathematik, Astronomie, und Physik* (Berlin, 1933), Abteilung B: Studien, II, p. 106.

³ Aristotle: *On Generation and Corruption*, Book I, Chap. ii, 316^a15–317^a17; A. Edel: *Aristotle's Theory of the Infinite* (New York: Columbia University Press; 1934), pp. 48–49, 76–78; T. L. Heath: *Mathematics in Aristotle* (Oxford: Oxford University Press; 1949), pp. 90, 117.

Pascal,⁶ Bolzano,⁷ Leibniz,⁸ Paul du Bois-Reymond,⁹ and Georg Cantor¹ to mention but a few. Writing on this issue more recently, P. W. Bridgman declared:

with regard to the paradoxes of Zeno . . . if I literally thought of a line as consisting of an assemblage of points of zero length and of an interval of time as the sum of moments without duration, paradox would then present itself.²

This criticism of the mathematical theory of physical space and time is a challenge to the basic Cantorean conceptions underlying analytic geometry and the mathematical theory of motion.³ Bridgman's view also calls into question such philosophies of science as rely on these conceptions for their interpretations of our mathematical knowledge of nature. Accordingly, it is essential that we inquire whether contemporary point-set theory succeeds in avoiding an inconsistency upon resolving positive linear intervals into extensionless point-elements.

In the present chapter, I shall endeavor to exhibit those features of present mathematical theory which do indeed preclude the existence of such an inconsistency. It will then be clear what kind of mathematical and philosophical theory does succeed in avoiding Zeno's mathematical (metrical) paradoxes of plurality, paradoxes that should be distinguished from his paradoxes of *motion*, on which I shall comment in Chapter Seven. My concern with the views which various writers have attributed to Zeno is exclusively systematic, and I make no claims whatever regarding

⁴ C. B. Boyer: *The Concepts of the Calculus* (New York: Hafner Publishing Company; 1949), p. 140.

⁵ *Ibid.*

⁶ *Ibid.*, p. 152.

⁷ *Ibid.*, p. 270, and B. Bolzano: *Paradoxes of the Infinite*, ed. D. A. Steele (New Haven: Yale University Press; 1951).

⁸ B. Russell: *The Philosophy of Leibniz* (London: G. Allen & Unwin, Ltd.; 1937), p. 114.

⁹ P. du Bois-Reymond: *Die Allgemeine Funktionentheorie* (Tübingen: Lauppische Buchhandlung; 1882), Vol. I, p. 66.

¹ G. Cantor: *Gesammelte Abhandlungen*, ed. E. Zermelo (Berlin: Julius Springer; 1932), pp. 275, 374.

² P. W. Bridgman: "Some Implications of Recent Points of View in Physics," *Revue Internationale de Philosophie*, Vol. III, No. 10 (1949), p. 490.

³ G. Cantor: *Gesammelte Abhandlungen*, *op. cit.*, p. 275.

the historicity of Zeno's arguments or concerning the authenticity of views which I shall associate with his name. According to S. Luria,⁴ Zeno invokes two basic axioms in his mathematical paradoxes of plurality. Having divided all magnitudes into positive and "dimensionless," i.e., unextended magnitudes, Zeno assumed that (1) the sum of an infinite number of equal positive magnitudes of arbitrary smallness must necessarily be infinite, (2) the sum of any finite or *infinite* number of "dimensionless" magnitudes must necessarily be *zero*.

The second of these axioms seems to command the assent of P. W. Bridgman and was also enunciated by the mathematician Paul du Bois-Reymond⁵ who then inferred that we cannot regard a line as an aggregate of "dimensionless" points, however dense an order we postulate for this aggregate. Zeno himself is presumed to have used these axioms as a basis for the following dilemma:⁶ If a line segment is resolved into an aggregate of infinitely many *like* elements, then two and only two cases are possible. Either these elements are of equal positive length and the aggregate of them is of infinite length (by Axiom 1) or the elements are of zero length and then their aggregate is necessarily of zero length (by Axiom 2). The first horn of this dilemma is valid but does not have relevance to the modern analytic geometry of space and time. It is the second horn that we must refute in the context of present mathematical theory if we are to solve the problem which we have posed.

To carry out this refutation, we must first ascertain the logical relationships between the modern concepts of metric, length, measure, and cardinality, when applied to (infinite) point-sets. For in the second horn of his dilemma, Zeno avers that a line cannot be regarded as an aggregate of points no matter what cardinality we postulate for the aggregate. And du Bois-Reymond endorsed this contention by reminding us that points are "dimensionless," i.e., unextended and by maintaining that if we conceive the line to be "merely an aggregate of points" then we are *eo*

⁴ S. Luria: "Die Infinitesimaltheorie der antiken Atomisten," *op. cit.*, p. 106.

⁵ P. du Bois-Reymond: *Die Allgemeine Funktionentheorie*, *op. cit.*, p. 66.

⁶ H. Hasse and H. Scholz: *Die Grundlagenkrise der griechischen Mathematik* (Charlottenburg: Pan-Verlag; 1928), p. 11.

ipso abandoning the view that "a line and a point are entirely different things."⁷

We see that du Bois-Reymond is conforming to the long intuitive tradition of using the concepts of length and dimensionality interchangeably to characterize (sensed) extension. It will therefore be best to begin our analysis by noting that we must distinguish the traditional *metrical* usage of the term "dimensionless" from the contemporary *topological* meaning of "zero dimension." This distinction has become necessary by virtue of the autonomous development of the topological theory of dimension apart from metrical geometry. Prior to this development, any positive interval of Cartesian n -space was simply called " n -dimensional" by definition. Thus line segments having length were called "one-dimensional" and surfaces having area "two-dimensional." By contrast, in the topological theory of dimension developed in the present century, it is a non-trivial *theorem* that lines are topologically one-dimensional, surfaces two-dimensional, and, generally, that Cartesian n -space is n -dimensional. In fact, it is this theorem which warrants the use of the name "dimension theory" for the branch of topology dealing with such non-metrical properties of point-sets as make for the validity of this theorem.⁸ By contrast, the traditional *metrical* sense of dimensionality identifies dimensionality with length or measure of extendedness. It is only the latter sense of "dimension" and "dimensionless" which is relevant to the metrical problem of this chapter. Hence I refer the reader to another publication of mine⁹ for an account of how the twentieth century theory of dimension can *consistently* affirm the following *additivity* properties for dimension in the *topological* sense of "zero-dimensional" and "one-dimensional": The point-set constituting the number-axis or any finite interval in it (e.g., an infinite straight line or a finite line segment respectively) is *one*-dimensional even though it is the set-theoretic sum of *zero*-dimensional subsets. The zero-dimensional subsets are: (a) Any unit point-set (such a set has a single point as its only member

⁷ P. du Bois-Reymond: *Die Allgemeine Funktionentheorie*, *op. cit.*, p. 65.

⁸ K. Menger: *Dimensionstheorie* (Leipzig: B. G. Teubner; 1928), p. 244.

⁹ A. Grünbaum: "A Consistent Conception of the Extended Linear Continuum as an Aggregate of Unextended Elements," *Philosophy of Science*, Vol. XIX (1952), pp. 290-95.

and hence can be loosely referred to as a "point," whenever such usage does not permit ambiguities), (b) Any *finite* collection of one or more points, (c) Any denumerable set (in particular the set of rational real points), (d) The set of irrational real points, which is non-denumerably infinite.

Accordingly, we must now deal with the *metrical* question of how the definition of length can *consistently* assign zero length to *unit* point-sets or individual points while assigning positive *finite* lengths to such unions (sums) of these unit point-sets as constitute a finite interval. To furnish an answer to the latter question will be to refute the second horn of Zeno's dilemma. We shall furnish an analysis satisfying these requirements after devoting some attention to the consideration of prior related problems.

Length or extension is defined as a property of point-sets rather than of individual points, and zero length is assigned to the *unit set*, i.e., to a set containing only a single point. While it is both logically correct and even of central importance to our problem that we treat a line interval of geometry as a set of point-elements, strictly speaking the definition of "length" renders it incorrect to refer to such an interval as an "aggregate of unextended points." For the property of being unextended characterizes unit point-sets but is not possessed by their respective individual point-elements, just as temperature is a property only of aggregates of molecules and not of individual molecules. The entities which can therefore be properly said to be unextended are *included in* but are not *members of* the aggregate of points constituting a line interval. Accordingly, the line interval is a union of unextended unit point-sets and, strictly, not an "aggregate of unextended points." Though strictly incorrect, I wish to use the latter designation in order to avoid the more cumbersome expression "union of unextended unit point-sets."

I shall now present such portions of the theory of metric spaces as bear immediately on our problem.

The structure characterizing the class of all real numbers (positive, negative, and zero) arranged in order of magnitude is that of a linear Cantorean continuum.¹

¹ E. V. Huntington: *The Continuum and Other Types of Serial Order* (2nd ed.; Cambridge: Harvard University Press; 1942), pp. 10, 44.

The Euclidean point-sets or "spaces" which we shall have occasion to consider are "metric" in the following complex sense:²

1. There is a one-one correspondence between the points of an n -dimensional Euclidean space E^n and a certain real coordinate system $(x_1, \dots, x_n)$
2. If the points x, y have the coordinates x_1, y_1 , then there is a real function $d(x, y)$ called their (Euclidean) *distance* given by

$$d(x, y) = \left\{ \sum_{i=1}^n (x_i - y_i)^2 \right\}^{1/2}.$$

The basic properties of this function are given by certain distance *axioms*.³

A finite interval on a straight line is the (ordered) set of all real points between (and sometimes including one or both of) two fixed points called the "endpoints" of the interval. Since the points constituting an interval satisfy condition 1 above in the definition of "metric," it is possible to define the "distance" between the fixed endpoints of a given interval. The *number* representing this distance is the *length* of the point-set constituting the interval. Let "a" and "b" denote respectively the points a and b or their respective real number coordinates, depending upon the context. We then define the length of a finite interval (a, b) as the non-negative quantity $b-a$ regardless of whether the interval $\{x\}$ is closed ($a \leq x \leq b$), open ($a < x < b$) or half-open ($a \leq x < b$ or $a < x \leq b$). (It is understood that the symbols "<" and "=" have a purely ordinal meaning here.) Therefore, the set-theoretic addition of a single point to an open interval (or to a half-open interval at the open end) has no effect at all on the *length* of the resulting interval as compared with the length of the original interval. In the limiting case of $a = b$, the interval is called "degenerate," and here the closed interval reduces to a set containing the single point $x = a$, while each of the other three intervals is empty. *It follows that the length of a degen-*

² S. Lefschetz: *Introduction to Topology* (Princeton: Princeton University Press; 1949), p. 28.

³ *Ibid.*

erate interval is zero. Loosely speaking, a single point has 0 length.⁴

This statement is clarified in item A of §10 in the Appendix

Zeno is challenging us to obtain a result differing from zero when determining the length of a finite interval on the basis of the known zero lengths of its degenerate subintervals, each of which has a single point as its only member. But since each positive interval has a *non-denumerable* infinity of degenerate subintervals, we see already that the result of determining the length of that interval by "compounding," in some unspecified way, the zero lengths of its degenerate subintervals is far less obvious than it must have seemed to Zeno, who did not distinguish between countably and non-countably infinite sets!

Although length is similar to cardinality in being a property of *sets* and not of the elements of these, it is essential to realize that the cardinality of an interval is not a function of the length of that interval. The independence of cardinality and length becomes demonstrable by combining our definition of length with Cantor's proof of the equivalence of the set of all real points between 0 and 1 with the set of all real points between *any* two fixed points on the number axis. It is therefore not the case that the longer of two positive intervals has "more" points. Once the independence of cardinality and length of such point-sets is established, it is possible to eliminate several of the confusions which have vitiated certain treatments of the infinite divisibility of intervals, as we shall see below. Thus, it will become impossible to infer in finitist manner that the division of an interval into two or more subintervals imparts to each of the resulting subintervals a cardinality lower than the cardinality of the original interval.

An interesting illustration of the independence of cardinality and length is provided by the so-called "ternary set" (Cantor discontinuum). This set has zero length (and zero dimension) while having the cardinality of the continuum.⁵

We shall be concerned with ascertaining why Zeno's *paradoxical* result that the length of a given *positive* interval (a, b) is

⁴ H. Cramér: *Mathematical Methods of Statistics* (Princeton: Princeton University Press; 1946), pp. 11, 19.

⁵ R. Courant and H. Robbins: *What is Mathematics?* (New York: Oxford University Press; 1941), p. 249.

zero is *not* deducible from the following two propositions in our geometry: (1) Any positive or non-degenerate interval is the union of a continuum of degenerate subintervals, and (2) The length of a degenerate (sub)interval is zero. It is obvious that *if the theory is consistent*, Zeno's result cannot be deducible. Such a result would contradict the proposition that the length of the interval (a, b) is $b-a$ ($a \neq b$). Furthermore, this result would be incompatible with Cantor's theorem that all positive intervals have the same cardinality regardless of length, for this theorem shows that no inference regarding the length of a non-degenerate interval can be drawn from propositions (1) and (2). In order to show later that our theory does have the required consistency, i.e., that it does not lend itself to the deduction of Zeno's paradoxical result, we must now consider the determination of (a) the length of the union of a *finite* number of non-overlapping intervals of known lengths, and (b) the length of the union of a *denumerable* infinity of such intervals.

If an interval "i" is the union of a *finite* number of intervals, no two of which have a common point, i.e., if

$$i = i_1 + i_2 + i_3 + \dots + i_n, \quad (i_p i_q = 0 \text{ for } p \neq q),$$

then it follows readily from the theory previously developed that the length of the total interval is equal to the *arithmetic sum* of the individual lengths of the subintervals. We therefore write

$$L(i) = L(i_1) + L(i_2) + L(i_3) + \dots + L(i_n).$$

If now we *define* the arithmetic sum of a progression of finite cardinal numbers as the limit of a sequence of partial arithmetic sums of members of the sequence, then a non-trivial proof can be given⁶ that the following theorem holds: The length of an interval which is subdivided into an enumerable number of subintervals without common points is equal to the arithmetic sum of the lengths of these subintervals.

Thus, both for a finite number and for a *countably* infinite number of non-overlapping subintervals, the length $L(i)$ of the total interval is an additive function of the interval i . The length of an interval is a numerical measure of the comprehensiveness

⁶ Cf. H. Cramér: *Mathematical Methods of Statistics*, *op. cit.*, pp. 19-21.

(extension) of that interval's membership but not of its cardinality. The latter does not depend upon the comprehensiveness of the membership of an interval.

It will be recalled that "length" was defined only for intervals. So far, we have not assigned any property akin to length to such sets as the set of all rational points between and including 0 and 1. There are many occasions, however, when it is desirable to have some kind of *measure* of the extensiveness, as it were, of point-sets quite different from intervals. Problems of this kind as well as problems encountered in the theory of (Lebesgue) integration have prompted the introduction of the generalized metrical concept of "measure" to deal with sets other than intervals as well. For details on the definition of "measure" for various kinds of point-sets, the reader is referred to Cramér⁷ and Halmos.⁸

Since the theory of infinite divisibility has been used fallaciously in an attempt to deduce Zeno's metrical paradox, we shall now point out the relevant fallacies before dealing with the crux of our problem to refute the second horn of Zeno's metrical dilemma.

In an exchange of views with Leibniz, Johann Bernoulli committed an important fallacy: he treated the actually infinite set of natural numbers as having a *last* or " ∞^{th} " term which can be "reached" in the manner in which an inductive cardinal can be reached by starting from zero.⁹ Bernoulli's view is clearly self-contradictory, since no discrete denumerable infinity of terms could possibly have a last term.

When giving arguments in behalf of his theory of infinitesimals, C. S. Peirce¹ committed the same Bernoullian fallacy by reasoning as follows: (1) the decimal expansion of an irrational number has an infinite number of terms, (2) the infinite decimal expansion has a *last* element at the "infinitieth place," and since the latter is "infinitely far out" in the decimal expansion, this element

⁷ H. Cramér: *Mathematical Methods of Statistics*, *op. cit.*, pp. 22ff.

⁸ P. R. Halmos: *Measure Theory* (New York: D. Van Nostrand Company, Inc., 1950.)

⁹ H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, p. 44.

¹ C. Hartshorne and P. Weiss (eds.): *The Collected Papers of Charles Sanders Peirce* (Cambridge: Harvard University Press, 1935), Vol. VI, para. 125.

is infinitely small or infinitesimal in comparison to finite magnitudes, (3) since continuity requires irrationals, continuity presupposes infinitesimals. Furthermore, the method of defining irrational points by nested intervals² was *misconstrued* by Du Bois-Reymond³ such that he was then able to charge it with committing the Bernoullian fallacy.⁴

We are now concerned with this fallacy, because it is always committed when the attempt is made to use the *infinite divisibility* of positive intervals as a basis for deducing Zeno's metrical paradox and for then denying that a positive interval can be an aggregate of unextended points. Precisely this kind of deduction of the paradox is attributed to Zeno by H. D. P. Lee⁵ and P. Tannery,⁶ both of whom seem to be unaware of the fallacy involved.

The following basic assumptions are involved in their version of Zeno's arguments:

1. Infinite *divisibility* guarantees the possibility of a *completable* process of "infinite division," i.e., of a completable yet infinite set of division operations.

2. The completion of this process of "infinite division" is achieved by the last operation in the series and terminates in

² R. Courant and H. Robbins: *What is Mathematics?*, *op. cit.*, pp. 68–69.

³ P. du Bois-Reymond: *Die Allgemeine Functionentheorie*, Vol. I, *op. cit.*, pp. 58–67.

⁴ Du Bois-Reymond's fundamental error lies in supposing that the method of nested intervals allows and requires the "coalescing" of the end-points of a supposedly "next-to-the-last" interval into a single point such that this "coalescing" is the *last step* in an infinite progression of nested interval formations. If the method in question did require such a coalescing, then it would indeed be as objectionable logically as is the Bernoullian conception of the ∞^{th} or last natural number. This is not the case, however, for while the method does indeed make reference to a progression of intervals, it neither allows nor requires that the irrational point is the "last" or " ∞^{th} " such "contracted" interval. Instead of appealing to "coalescence," the method specifies the irrational point *by the mode of variation* of the intervals in the *entire sequence*. It is therefore a property of the entire sequence which enables us to define the kind of point which is being asserted to exist. It would seem that Du Bois-Reymond permitted himself to be misled by pictorial language like "the interval contracts into a point."

⁵ H. D. P. Lee: *Zeno of Elea* (Cambridge: Luzac and Company; 1936), p. 23.

⁶ P. Tannery: "Le Concept Scientifique du Continu: Zénon d'Élée et Georg Cantor," *Revue Philosophique*, Vol. XX, No. 2 (1885), pp. 391–92.

saw that "infinite *division*" of this kind is self-contradictory. Therefore, confusion is avoided by basing our analysis only on infinite *divisibility* and on the actual infinite "dividedness" of sets.⁸

It follows from our definition of division and from the properties of finite sets that the division of a *finite* point-set of two or more members necessarily effects a reduction in its cardinality. This reduction is in marked contrast to the behavior of *intervals* since their cardinality remains unaffected by division. Since the degenerate interval has no proper non-empty subset, that unique kind of interval is *indivisible*. We see that on our theory, *indivisibility is not a metrical property* at all but a set-theoretic one. This theory has enabled us to assign a precise meaning to the indivisibility of a unit point-set by (1) defining division as an operation on sets only and not on their elements, (2) defining divisibility of finite sets as the formation of proper non-empty subsets of these, (3) showing that the degenerate interval is indivisible by virtue of its lack of a subset of the required kind.

It is of importance to realize that our analysis has shown how we can assert the following *two* propositions *perfectly consistently*:

1. The line and intervals in it are *infinitely divisible*.
2. The line and intervals in it are each a union of *indivisible* degenerate intervals.

We are now prepared to deal with the crux of our problem by using point-set theory to refute the second horn of Zeno's metrical dilemma.

Since a positive interval is the union of a continuum of degenerate intervals,⁹ we must now determine what meaning, *if any*, we can assign to "summing" the lengths of these degenerate intervals so as to obtain the length of the total interval. The answer which we shall give to this problem will not be *ad hoc*, since the reasoning on which it is based will not depend upon

⁸ Nevertheless it is often convenient *by way of elliptic parlance* to designate the membership of a set through mention of an actual infinity of operations which, as it were, "generate" the elements of the set in question.

⁹ The word "continuum" can designate either the ordering structure of the real numbers or their cardinality. The context will indicate which of these meanings is intended or whether both are jointly involved.

the particular lengths which Zenonians wish us to "compound" but rather on the fact that the number of lengths to be "added" is *not denumerable*.

Earlier, we determined the length of the union of a *finite* number of non-overlapping intervals of known lengths on the basis of these latter lengths. In addition, we made a corresponding determination of the length of the union of a *denumerable* infinity of non-overlapping intervals. If now we attempt to subdivide an interval into a *non-denumerable* infinity of such intervals, we find that they *cannot be non-degenerate*. For Cantor has shown that any collection of positive non-overlapping intervals on a line is at most denumerably infinite.¹ It follows that the degenerate subintervals which are at the focus of our interest are the only kind of (non-overlapping) subintervals of which there are non-denumerably many in a given interval. Quite naturally, therefore, they create a special situation. The latter is due to the fact that our theory does *not* assign any meaning to "forming the *arithmetic* sum," when we are attempting to "sum" a *super-denumerable* infinity of individual numbers (lengths)! This fact is independent of whether the individual numbers in such a non-denumerable set of numbers are zeros or finite cardinal numbers differing from zero.

Consequently, the theory under discussion cannot be deemed to be *ad hoc* for precluding the possibility of "adding," in Zenonian fashion, the zero lengths of the continuum of points which "compose" the interval (a, b) to obtain zero as the length of this interval. Though the finite interval (a, b) is the union of a continuum of degenerate subintervals, *we cannot meaningfully determine its length in our theory by "adding" the individual zero lengths of the degenerate subintervals*. We are here confronted with an instance in which set-theoretic addition is possible while arithmetic addition is not.

We have shown that geometrical theory as here presented does *not* have the paradoxical feature of both assigning the non-zero length $b-a$ to the interval (a, b) *and* permitting the inference that (a, b) must have zero length on the grounds that its points each have zero length. More precisely, we have shown that geo-

¹ G. Cantor: *Gesammelte Abhandlungen*, *op. cit.*, p. 153.

metrical theory can consistently affirm the following four propositions simultaneously:

1. The finite interval (a, b) is the union of a continuum of degenerate subintervals.
2. The length of each degenerate (sub)interval is 0.
3. The length of the interval (a, b) is given by the number $b-a$.
4. The length of an interval is not a function of its cardinality.

Our analysis has manifestly refuted the Zenonian allégation of inconsistency if made against set-theoretical geometry.

The set-theoretical analysis of the various issues raised or suggested by Zeno's paradoxes of plurality has enabled me to give a *consistent* metrical account of an extended line segment as an aggregate of unextended points. Thus, Zeno's mathematical paradoxes are avoided in the formal part of a geometry built on Cantorean foundations. The consistency of the metrical analysis which I have given depends crucially on the *non-denumerability* of the infinite point-sets constituting the intervals on the line. For the length (measure) of an *enumerably* infinite point-set (like the set of rational points between and including 0 and 1) is *zero* (upon denumeration of the set), as can readily be inferred from our analysis above.² Thus, if any set of rational points were regarded as constituting an extended line segment, then the customary mathematical theory under consideration could assert the length of that segment to be greater than zero only at the cost of permitting itself to become self-contradictory!

It might seem that my conclusion concerning the fundamental logical importance of non-denumerability could be criticized in the following way: The need for non-denumerably infinite point-sets to avoid metrical contradictions derives from defining the arithmetic sum of an infinite series of numbers as the limit of its partial sums as we have done.³ Without this definition, it would not have been possible to infer that the sum of the individual zero lengths of the points (unit point-sets) in an enumerable point-set turns out to be *zero* upon denumeration of that set. Consequently, by omitting this definition, it would presum-

² Cf. also H. Cramér: *Mathematical Methods of Statistics*, *op. cit.*, p. 257.

³ Cf. also G. H. Hardy: *A Course of Pure Mathematics* (9th ed.; New York: The Macmillan Company; 1945), pp. 145-47.

ably have been possible to assign a finite length to certain enumerable sets without contradiction. Thus it might be argued that a non-denumerably infinite point-set is indispensable for consistency only *relatively* to a formulation of the theory containing the definition that the arithmetic "sum" of an infinite series of numbers is the limit of its partial sums.

My reply to this objection is that the omission of the latter definition from the system would entail incurring the loss of the theory of infinite convergent series in analysis and geometry and of whatever part of the theory depends in part or whole upon the presence of this definition in the foundations. It follows that instead of being a merely incidental feature of the theory, the introduction of the definition in question was dictated by important theoretical considerations. The requirement that the points on the line be non-denumerably infinite which must then be satisfied to insure metrical consistency therefore has a corresponding significance.

Consequently, it is not clear how theorists whose philosophical commitments do not allow them to avail themselves of a super-denumerable set would avoid Zeno's mathematical paradoxes in keeping with their theory.

Proponents of Zeno's view might still argue that this *arithmetical* rebuttal is unconvincing on purely geometric grounds, maintaining that if extension (space) is to be composed of elements, these must themselves be extended. Specifically, geometers like Veronese objected⁴ to Cantor that in the array of points on the line, their extensions are all, as it were, "summed geometrically" before us. And from this geometric perspective, it is not cogent, in their view, to suppose that even a super-denumerable infinity of unextended points would be able to sustain a positive interval, especially since the Cantorean theory can claim arithmetical consistency here only because of the obscurities that obligingly surround the meaning of the arithmetic "sum" of a super-denumerable infinity of numbers.

Is this objection to Cantor conclusive? I think not. Whence does it derive its plausibility? It would seem that it achieves persuasiveness via a tacit appeal to a *pictorial* representation of the points of mathematical physics in which they are arrayed in

⁴ See E. W. Hobson: *The Theory of Functions of a Real Variable* (2nd ed.; Cambridge: Cambridge University Press, 1921) Vol. I, pp. 56-57.

the consecutive manner of beads on a string to form a line. But the properties that any such representation imaginatively attributes to points are not even allowed, let alone prescribed, by the formal postulates of geometric theory. The spuriousness of the difficulties adduced against the Cantorean conception of the line becomes apparent upon noting that not only the cardinality of its constituent points altogether eludes pictorialization but also their dense ordering: between any two points, there is an infinitude of others. Thus, in complete contrast to the discrete order of the beads on a string, *no* point is immediately adjacent to any other.

These considerations show that from a genuinely geometric point of view, a physical interpretation of the formal postulates of geometry cannot be obtained by the inevitably misleading pictorialization of *individual* points of the theory. Instead, we can provide a physical interpretation quite unencumbered by the intrusion of the irrelevancies of *visual* space, if we associate *not* the term "*point*" but the term "*linear continuum of points*" of our theory with an appropriate body in nature. By a point of this body we then mean nothing more or less than an element of it possessing the formal properties prescribed for points by the postulates of geometry. And, on this interpretation, the ground is then cut from under the geometric *parti pris* against Cantor by the modern legatees of Zeno.

I can best illustrate the importance of being mindful of the need for a super-denumerable point-set to avoid Zeno's metrical contradictions by showing how (1) Russell, and (2) the proponents of a strict operationism in geometry were led into error by failing to be aware of this need.

1. Russell neglected the essential contribution made by the cardinality and ordinal structure of the continuum toward the avoidance of Zeno's mathematical paradoxes in the mathematical theory of motion. This philosophical neglect of his is clear in the following passages:

Mathematicians have distinguished different degrees of continuity, and have confined the word "continuous," for technical purposes, to series having a certain high degree of continuity. But for *philosophical* purposes, all that is important in continuity is introduced by the lowest degree of continuity, which is called "compactness" [i.e., denseness].

... What do we mean by saying that the motion is continuous? It is not necessary for our purposes to consider the whole of what the mathematician means by this statement: *only part of what he means is philosophically important*. One part of what he means is that, if we consider any two positions of the speck occupied at any two instants, there will be other intermediate positions occupied at intermediate instants. . . .⁵

We know that the mere existence of the denseness property guarantees only a denumerably infinite point-set. Since a superdenumerably infinite point-set is required by the demands of metrical consistency, it follows that there are *philosophical* reasons for requiring a higher degree of continuity than is insured by the denseness property alone.

My analysis has shown that within the framework of the mathematics used in physical theory we cannot postulate the line to consist of the rational points alone, since the latter constitute a *denumerable* set. The choice between postulating the line to consist of the real points and resolving it into the rational points alone is therefore *not* a matter of mere arithmetical convenience in the analytic part of geometry or of mere aesthetic satisfaction to the scientist. Russell is unaware of this fact, for he asserts that a "rational space" may be actual. Says he: "a space . . . in which the points of a line form a series ordinally similar to the rationals, will, with suitable axioms, be empirically indistinguishable from a continuous space, *and may be actual*."⁶

2. The Greeks certainly were *not* led to the discovery of incommensurable magnitudes by merely *operationally* carrying out

⁵ B. Russell: *Our Knowledge of the External World* (London: George Allen & Unwin, Ltd.; 1926), pp. 138, 139-40; my italics.

⁶ B. Russell: *The Principles of Mathematics* (Cambridge: Cambridge University Press; 1903), p. 444; my italics. A similar criticism applies to Dedekind. He maintains that if we postulated a discontinuous space consisting of the algebraic points alone, then "the discontinuity of this space would not be noticed in Euclid's science, would not be felt at all" (R. Dedekind: *Essays on the Theory of Number* [Chicago: Open Court Publishing Company; 1901], pp. 37-38). Since the set of algebraic points is still denumerable (A. Fraenkel: *Einleitung in die Mengenlehre* [New York: Dover Publications, Inc.; 1946], p. 40), the length (measure) of a segment consisting entirely of such points is paradoxically both zero and positive. I therefore cannot follow F. Waismann, who comments approvingly on Dedekind's statement by saying: "As to physical space one has become accustomed to conceding the justification of this concept" (F. Waismann: *Introduction to Mathematical Thinking* [New York: F. Ungar Publishing Company; 1951], p. 212).

the iterative transport of measuring sticks.⁷ And it is impossible to show by direct physical operations alone that there are hypotenuses whose length cannot be represented by any rational number. For the denseness of the rational points guarantees that we can never claim anything but a rational result on the strength of operational accuracy alone. A radical operationist approach to geometry might therefore suggest that this science be constructed so as to use only the system of rational points.⁸ The analysis given in this chapter has shown that in the absence of an alternative mathematical theory which is demonstrably viable for the purposes of physics, such an operationist approach to geometry and to the theoretical measurables of physics must be rejected on logical grounds alone.

Though I am claiming to have shown that set-theoretical geometry can successfully meet the challenge of Zeno's mathematical paradoxes *if it is otherwise consistent*, I am *not* asserting that Zeno and his followers were either mistaken within the context of the mathematical knowledge of their time or that they were deficient in philosophical acumen.

⁷ For the historical details, see K. von Fritz: "The Discovery of Incommensurability by Hippasus of Metapontum," *Annals of Mathematics*, Vol. XLVI (1945).

⁸ Cf. the approximative geometry of J. Hjelmslev (J. Hjelmslev: "Die natürliche Geometrie," *Abhandlungen aus dem mathematischen Seminar der Hamburger Universität*, Vol. II (1923), p. 1 ff.) and Weyl's comments on it (H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, pp. 143-44).

PART II

Philosophical Problems of the Topology of Time and Space

See Appendix §11

Chapter 7

THE CAUSAL THEORY OF TIME

See
Append §§11 and 12

The causal theory of time, which had occupied an important place in the thought of Leibniz and of Kant, again became a subject of central philosophic interest during the current century after its detailed elaboration and logical refinement at the hands of G. Lechalas,¹ H. Reichenbach,² K. Lewin,³ R. Carnap,⁴ and H. Mehlberg.⁵ Specifically, it earned its new prominence in recent decades by its role in the magisterial and beautiful constructions of the relativistic topology of both time and space by Reichenbach⁶ and Carnap.⁷ More recently, I used the causal

¹ G. Lechalas: *Etude sur l'espace et le temps* (Paris: Alcan Publishing Company; 1896).

² H. Reichenbach: *Axiomatik der relativistischen Raum-Zeit-Lehre* (Braunschweig: Friedrich Vieweg & Sons; 1924).

³ K. Lewin: "Die zeitliche Geneseordnung," *Zeitschrift für Physik*, Vol. XIII (1923), p. 16.

⁴ R. Carnap: "Über die Abhängigkeit der Eigenschaften des Raumes von denen der Zeit," *Kantstudien*, Vol. XXX (1925), p. 331.

⁵ H. Mehlberg: "Essai sur la théorie causale du temps," *Studia Philosophica*, Vol. I (1935), and Vol. II (1937).

⁶ H. Reichenbach: *Philosophie der Raum-Zeit-Lehre* (Berlin: Walter de Gruyter & Company; 1928), esp. pp. 307–308. The English translation entitled *The Philosophy of Space and Time*, *op. cit.*, will hereafter be cited as "PST."

⁷ R. Carnap: *Abriss der Logistik* (Vienna: Julius Springer: 1929), Sec. 36, pp. 80–85. Cf. also his *Symbolische Logik* (Vienna: Julius Springer; 1954), Secs. 48–50, pp. 169–81; an English translation, *Introduction to Symbolic Logic and Its Applications* was published by Dover Publications, Inc., New York, in 1958. For an interesting comparison of Kant's version of the theory with the conceptions propounded by Carnap, Reichenbach, and Mehlberg, see H. Scholz: "Eine Topologie der Zeit im Kantischen Sinne," *Dialectica*, Vol. IX (1955), p. 66.

theory of time to show semantically that, with respect to the relation "*later than*," the events of physics can meaningfully possess the seemingly counter-intuitive *denseness* property of the linear Cantorean continuum. And, in this way, I was able to supply the semantical *nervus probandi* which had been lacking in Russell's mathematical refutation of Zeno's paradoxes of motion.⁸

In order to assess the causal theory of time philosophically, we must inquire whether the physical meaning of the primitive *asymmetric* causal relation employed in most versions of it can be understood without possessing a *prior* understanding of the *temporal* terms which it is intended to define. I shall therefore devote the present chapter to an examination of the major defenses and criticisms of the causal theory of time, beginning with the Reichenbachian version of the theory, which is based on the mark method.⁹

My aim will be to show that (1) Reichenbach's mark method fails to define a serial temporal order within the class of physical events, being vitiated by circularity in its attempt to define the required asymmetric relation; (2) although the classical Leibniz-Reichenbach version of the causal theory of time is vulnerable to these criticisms, it is possible to define *temporal betweenness* on the basis of the postulate of causal continuity for reversible mechanical processes; but nomologically contingent boundary conditions must prevail, if the betweenness defined by causal continuity is to have the formal properties of the triadic relation ordering the points on an *open* (straight) *undirected* line, hereafter called "o-betweenness"; alternatively, the boundary conditions may issue in a temporal order exhibiting the formal properties possessed by the order of points on a *closed* (circular) *undirected* line with respect to a tetradic relation of separation, hereafter called "separation closure"; (3) as distinct from Reichenbach's mark method, a significant *modification* of his most recent account of the *anisotropy* of time as a statistical

⁸ A. Grünbaum: "Relativity and the Atomicity of Becoming," *op. cit.*, esp. pp. 168-69. Much material concerning time in this article, especially the endorsement of Reichenbach's version of the causal theory of time, is superseded by the treatment given in this book. In addition, the article contains distorting misprints.

⁹ Cf. H. Reichenbach: *PST*, *op. cit.*, pp. 135-38.

property of the *entropic* behavior of *space ensembles* of "branch systems,"¹ is successful; but (4) his conception of becoming as the transience of a physically-distinguished "now" along one of the two opposite directions of time is untenable. Of these four contentions, only (1) and (2) will be defended in the present chapter, (3) and (4) being deferred to Chapter Eight (Part A, II) and Chapter Ten, respectively.

Reichenbach introduces his mark method by giving the following topological coordinative definition of temporal sequence:² "If E_2 is the effect of E_1 , then E_2 is said to be later than E_1 ." To show that causality defines an asymmetric temporal relation without circularity, he invites attention to the fact that when E_1 is the cause of E_2 , small variations in E_1 in the form of the addition of a marking event e to E_1 will be connected with corresponding variations in E_2 , but not conversely. Thus, if we denote an event E that is slightly varied (marked) by " E^e ," we shall find, he tells us, that we observe only the combinations

$$E_1E_2 \quad E^e_1E^e_2 \quad E_1E^e_2$$

but *never* the combination

$$E^e_1E_2.$$

In the observed combinations, the events E_1 and E_2 play an asymmetric role, thereby defining an order, and it is clear that this order would be unaffected by interchanging the subscripts in the symbols naming the events involved. The event whose name does not have an "e" in the non-occurring combination is called the effect and the later event.³

¹ H. Reichenbach: "Les fondements logiques de la mécanique des quanta," *Annales de l'Institut Henri Poincaré*, Vol. XIII (1953), pp. 140-58, and "La signification philosophique du dualisme ondes-corpuscules," in A. George (ed.) *Louis de Broglie, Physicien et Penseur* (Paris: Editions Albin Michel; 1953), pp. 126-34. The most detailed account is given in his book *The Direction of Time* (Berkeley: University of California Press; 1956). Hereafter this book will be cited as "DT."

² H. Reichenbach: *PST*, *op. cit.*, p. 136. Cf. also G. W. Leibniz: "Initia rerum mathematicorum metaphysica," *Mathematische Schriften*, ed. Gerhardt (Berlin: Schmidt's Verlag; 1863), Vol. VII, p. 18; H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, pp. 101, 204, and "50 Jahre Relativitätstheorie," *Naturwissenschaften*, Vol. XXXVIII (1951), p. 74; H. Poincaré: *Letzte Gedanken*, trans. Lichtenecker (Leipzig: Akademische Verlagsgesellschaft; 1913), p. 54.

³ H. Reichenbach: *PST*, *op. cit.*, p. 137.

Reichenbach's formulation of his principle contains no restriction to causal chains which are either materially genidentical, such as stones, or possess the quasi-material genidentity of individual light rays. Thus, although the illustrations he gives of the principle do involve stones and light rays, it has been widely interpreted as not requiring this kind of restriction. Moreover, it is not clear even from his illustration of the stone whether the members of his three observed pairs of events are to belong to *three* rather than to only one or two different genidentical causal chains, a specification which is of decisive importance if such a restriction is to obviate certain of the criticisms of his method which are about to follow. Accordingly, those objections which might thus have been obviated will also be stated.

An illustration of the use of Reichenbach's method given by W. B. Taylor in an attempt to demonstrate its independence of prior knowledge of temporal order⁴ will serve to let me set forth the objections to it.

An otherwise dark room has two holes in opposite walls such that a single light ray traverses the room. Since the light source is hidden in the wall behind one of the two holes, we are not able to tell, as we face the light ray from a perpendicular direction, whether it travels across the room from left to right or right to left and thus do not know the location of the light source. In order to ascertain the direction of this causal process, let us isolate three events E_L in it which occur at the left end of the beam and also three events E_R of the process occurring at the right end. And now suppose that *one* of the three events at *each* end is *marked*, say, by means of the presence of transparent colored glass at the point and instant of its occurrence. The crux of the matter lies in Reichenbach's claim that if the events E_L are the respective (partial) causes of the events E_R , we shall observe only the combinations

$$E_L E_R \quad E_L^{\circ} E_R^{\circ} \quad E_L E_R^{\circ}$$

and *never* the combination

$$E_L^{\circ} E_R.$$

⁴ W. B. Taylor: *The Meaning of Time in Science and Daily Life*, doctoral dissertation (Los Angeles: University of California at Los Angeles, 1953), pp. 37-39.

We shall now see that in order to obtain these particular kinds and pairs of events, on which the method relies to define an asymmetric relation, Reichenbach must either make tacit and inadmissible use of prior temporal knowledge or invoke the special requirement of *irreversible* marking processes. For in the absence of information concerning the temporal order within either triplet of events at the left end or at the right end, all that we can say is that our observations at the two ends can be represented by *temporally neutral* triangular arrays:

$$\begin{array}{cc} E_L & E_R^* \\ E_L & E_L^* & E_R & E_R^* \end{array}$$

And, again not presupposing temporal information, we are entitled to *interpret* our data to the effect that we observed the three combinations

$$E_L E_R^* \quad E_L E_R^* \quad E_L^* E_R$$

but *not* $E_L^* E_R^*$.

In the context of Reichenbach's program, this latter interpretation seems to be fully as legitimate as the formation of his own particular grouping, which likewise constitutes a mere interpretation. But the legitimacy of this alternate interpretation is fatal to Reichenbach's claim that E_L and E_R play an asymmetric role, since the alternate interpretation contains the combination $E_L^* E_R$, which is the very combination that he had to rule out in order to show that E_R is later than E_L !

Even if we know the temporal sequence *within* each of the two triplets of events to be, say,

$$\begin{array}{cc} E_L & E_R \\ E_L^* & \text{and} & E_R^* \\ E_L & & E_R^* \end{array}$$

where the *downward* direction is the direction of increasing time, it is still not clear which particular event at the right end is to be associated with a particular event at the left end to form a pair. Hence, Reichenbach's causal theory allows us to form event-pairs so as to obtain both his own asymmetric interpretation *and* the alternate interpretation above even from the two

internally *ordered* triplets of events. The alternate interpretation is obtainable by combining the first event in the E_L column with the second in the E_R column, the respective third events in the two columns and the second in the E_L column with the first in the E_R column.

Nor is this legitimate alternate interpretation the only source of difficulties for the mark-method. For without a criterion for uniting the spatially separated events under consideration in pairs, we might unwittingly mark the pulse that actually emanated from E_L^* upon its arrival at the right, instead of marking one of the other two pulses, which emanate from *unmarked* events at the left. Since the pulse coming from E_L^* will already be bearing a mark, the marking of that particular pulse upon arrival at the right will be *redundant*. We could then interpret our data as forming the combinations

$$E_L E_R \quad E_L^* E_R^* \quad E_L E_R$$

in which E_L and E_R occur with complete *symmetry*.⁵ The proponent of the mark method cannot, of course, avert this embarrassing consequence by requiring that we mark at the right end only a pulse not already bearing a mark from before.⁶

Equally unavailing is the following attempt by W. B. Taylor to justify the mark method:

In the example concerning the light ray, it was said that we first mark E_L and *then* (a moment later) see if the mark appears on E_R . As it stands, this procedure employs time order, which would be undesirable for the purposes at hand. But this way of stating the procedure can be eliminated by saying that we mark E_L and observe (tenseless form of the verb "observe")

⁵ Using the term "partial cause" so as to allow some of its designata to be non-simultaneous, Hugo Bergmann offers another argument for the variational *symmetry* of cause and effect (*Der Kampf um das Kausalgesetz in der jüngsten Physik* [Braunschweig: F. Vieweg and Sons: 1929], pp. 16-19) by showing that Reichenbach succeeds merely in proving that the variation of one of the *partial* causes does not vary the other, but *not* that the effect can be varied without varying the cause. Reichenbach's retort ("Das Kausalproblem in der Physik," *Naturwissenschaften*, Vol. XIX [1931], p. 719) to Bergmann is superseded by his subsequent acknowledgement (DT, *op. cit.*, pp. 198-99) of an inadequacy here which we shall soon consider.

⁶ For a similar objection, cf. H. Mehlberg: "Essai sur la théorie causale du temps," *op. cit.*, Vol. I, pp. 214-15.

whether that same mark appears (again tenseless) on the E_R which is causally connected with E_L . The reason that time order appears to be presupposed is perhaps that we must use verbs (e.g., observe, appear) to describe the procedure of marking and because English verbs in their usual usage are in a tensed, and hence time-referent form. If the observer does not himself mark the events, but instead relies on other agencies to provide the marks, then this apparent circularity does not arise. He, in this case, simply observes the event pairs $E_L E_R$ to see in what arrangement the mark e appears.⁷

This argument will not do, since we saw that it is not possible, without knowledge of the temporal order, to "simply observe" particular, observationally *given* event-pairs $E_L E_R$ to see in what arrangement the mark e appears. The issue is not that time is presupposed in having to mention in temporal succession the names of the members of already-given event-pairs or in stating first the result of our observation on one event in such a pair and then on the other. Instead, the difficulty is that time order is presupposed in assuring and singling out the membership of Reichenbach's specific three event-pairs to begin with, though not in the internal arrangement of their members after the pairs themselves are already chosen. Once we grant the uniqueness of the Reichenbachian choice of event-pairs, it is quite true that temporal order is only apparently presupposed in the description of the experiment, since E_L and E_R do occur asymmetrically in those pairs, independently of the order in which they are named within the pairs.

A further difficulty, which will turn out to have an **important** bearing on the relation of causality to the anisotropy of time, lies in the fact that in the case of *reversible* marking processes (if there be such), the mark method must make illicit use of temporal betweenness to preclude failure of its experiments. For, as Mehlberg has rightly observed in his searching paper on the causal theory of time,⁸ if the mark e were removed, in some way or other, from a signal originating at E_L^0 *while that signal is in transit*, the experiment would yield the combination $E_L^0 E_R$, which is precisely the one disallowed by Reichenbach. To prevent such

⁷ W. B. Taylor: *The Meaning of Time in Science and Daily Life*, *op. cit.*, pp. 38–39.

⁸ H. Mehlberg: "Essai sur la théorie causale du temps," *op. cit.*, Vol. I, p. 214; also pp. 207–57.

an eventuality, the mark-method must either incur failure by requiring that the physical system under consideration be *closed* during the *time interval between* E_L^a and E_R^a ,⁹ or it must have recourse to an irreversible marking process such as passing white light through a color filter.

The question arises, therefore, whether the requirement that the marking processes be *irreversible* does not constitute an invocation of a new criterion of temporal order, since it is based on a much more restricted class of kinds of occurrences than the one to which causality is held to apply, the latter including *reversible* processes. Indeed, if the meaning of causality is correctly explicated by the mark method, and if that method—its other difficulties aside—can hope to be successful in defining a serial temporal order via an *asymmetric* causal relation for only those causal processes which are irreversible or which are rendered irreversible by its application, then two conclusions follow: first, causality as such is *not sufficient* to define those topological properties which are conferred on physical time by irreversible processes, and second, Reichenbach's causal criterion cannot be logically *independent* of a criterion grounded on such processes. Though declining to discuss the issue in his book of 1924, Reichenbach admitted then that the independence of these two criteria is open to question while nevertheless affirming the autonomy of the causal criterion and characterizing the concordance of the temporal orders based on causality, on the one hand, and irreversibility, on the other, as an empirical fact.¹ In his last paper on the subject, published just before his death, and in his

⁹ In another connection, Reichenbach defines a closed system as a system "not subject to differential forces" (*PST*, *op. cit.*, pp. 22-23 and 118), differential forces being forces whose presence is correlated with *changes* of varying degree in different kinds of materials. But the absence of a change at any given instant t means the constancy of a certain value (or values) *between* the termini of a time interval containing the instant t (no *anisotropy* of time being assumed). Thus Reichenbach recognizes that the concept of closed system presupposes the ordinal concept of temporal betweenness. We see incidentally that the meaning of temporal betweenness is presupposed by the statement of the second law of thermodynamics, which concerns *closed* systems.

¹ H. Reichenbach: *Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, pp. 21-22. For critical comments on Reichenbach's early views concerning this issue, see E. Zilsel: "Über die Asymmetrie der Kausalität und die Einsinnigkeit der Zeit," *Naturwissenschaften*, Vol. XV (1927), p. 282.

posthumous book *DT*, he abandoned the ambitious program of the mark method to define an anisotropic serial time at one stroke. Instead, he offered a construction in which he relied on a certain kind of thermodynamic irreversibility and not merely on causality in an attempt to achieve this purpose.² In the next chapter we shall examine this construction critically and see in what specific sense irreversible processes can be held to define the anisotropy of time.

Before proceeding to a consideration of other issues, a problem encountered by the causal theory of time merits being stated.

On the causal theory of time, the existence of an *actual* causal chain linking two events is only a sufficient and not a necessary condition for their sustaining a relation of being temporally apart. Such a relation is likewise held to obtain between any two events for which it is merely physically *possible* to be the termini of a causal chain even if no actual chain connects them. Thus, causal connectibility rather than actual causal connectedness is the defining relation of being temporally apart. And causal non-connectibility rather than non-connectedness is the defining relation of topological simultaneity. But as Mehlberg has noted,³ physical possibility, in turn, must then be definable or understood in such a manner as not to presuppose the ordinal concepts of time which enter into the laws that tell us what causal processes are physically possible. It is not clear that Mehlberg's theory of causal decompositions⁴ provides a basis on which the difficulty might be circumvented. For he maintains that *not-E* must be held to be a physical event if *E* is such an event. And on the strength of this contention, Mehlberg thinks that any two events which are the termini of physically possible causal chains must therefore be held to be actually causally connected. Mehlberg thus proposes to guarantee the actual rather than merely poten-

² H. Reichenbach: "Les fondements logiques de la mécanique des quanta," *op. cit.*, pp. 137-38; *DT*, *op. cit.*, p. 198n. For his earlier views, see *PST*, *op. cit.*, pp. 138-39; "The Philosophical Significance of the Theory of Relativity," *op. cit.*, pp. 304-306, and "Ziele und Wege der physikalischen Erkenntnis," *Handbuch der Physik* (Berlin: Julius Springer; 1929), Vol. IV, pp. 53, 59-60, 64, 65.

³ H. Mehlberg: "Essai sur la théorie causale du temps," *op. cit.*, Vol. I, pp. 191, 195, and Vol. II, p. 143.

⁴ *Ibid.*, Vol. I, pp. 165-66, 240-41, and Vol. II, pp. 145-46, 169-72.

tial existence of all the required causal chains by asserting that for any event not belonging to a class of simultaneous events, there is at least one event in that class with which the first event is causally connected.

We saw that the mark method is vitiated by circularity, because it explicates the causal relation in such a manner that: first, the method is then required to provide a criterion for designating, within a pair of causally connected events, the *one* event which is *the* cause of the other by virtue of being the earlier of the two in anisotropic time, and second, the method leaves itself vulnerable to the charge of having tacitly employed temporal criteria to secure the particular pairs of events which are essential to its success. Despite the failure of the variational conception of causality offered by the mark method, the kind of causality exhibited by reversible processes is competent, as we shall see shortly, to define some of the topological features of time. Although Leibniz and more recent proponents of his causal definition of the relation "later than" were mistaken concerning the character and extent of the logical connection between the topology of time and the structure of causal chains, their affirmation of the existence of such a connection was sound. We shall see now that reversible genidentical causal processes do indeed define an order of temporal *betweenness* (albeit only in part) and also relations of *simultaneity* in the class of physical events.

By contrast to the mark method, our construction will have the following features: First, we consider a kind of causal relation between two different events which is *symmetric* and involves no reference at all to one of the two events being *the* cause of the other, the criterion for the latter characterization to be supplied subsequently in Chapter Eight; second, we shall eschew resting the asymmetry, and thereby, the seriality of the relation of "earlier than" on the causal relation itself; in fact, the latter relation will be seen to be neutral with respect to whether the

⁵ *Ibid.*, Vol. II, p. 169.

⁶ For a discussion of the ancestral role of the Leibnizian conception, see H. Reichenbach: "Die Bewegungslehre bei Newton, Leibniz und Huyghens," *Kantstudien*, Vol. XXIX (1924), p. 416, and the English translation of this paper by Maria Reichenbach in the posthumous volume of Reichenbach's selected essays entitled *Modern Philosophy of Science* (London: Routledge and Kegan Paul; 1959), pp. 46-66.

temporal order has the formal properties of the "o-betweenness" exhibited by the points on an undirected straight line or those of "separation-closure" relating the points on an undirected circle;⁷ instead of being made to depend on causality itself, the o-betweenness of time will depend for its existence on the *boundary* conditions, which determine the relations of the various causal chains to one another; and third, instead of construing causality variationally in the manner of the mark method, we shall begin with material objects each of which possesses genidentity (i.e., the kind of sameness that arises from the persistence of an object for a period of time) and whose behavior therefore provides us with genidentical causal chains. And we shall then consider the causal relation between any pair of events belonging to a genidentical causal chain.

Consider any (ideally) *reversible* genidentical causal process such as the rolling of a ball on the floor of a room along a path connecting points P and P' of the floor. So long as we confine ourselves to processes which are both nomologically and de facto reversible, we forsake all reliance on the anisotropy of time, since the latter depends—as we shall see in Chapter Eight—on processes which are de facto (and statistically) *irreversible*. Hence in our present temporally *isotropic* context, which excludes every kind of irreversible process, there is no objective physical basis for singling out *one* of two causally connected events as "the" cause and the other as "the" effect. For the designation of one of these events as "the" cause depends on its being the *earlier* of the two, and, in the absence of irreversible processes, the latter characterization expresses not an objective physical relation sus-

See Append. §17 for a discussion of the rival conception of retro-causation

⁷ The order of points on an undirected circle which we have here called "separation-closure" is generally called "separation of point pairs" in the literature. It is the order of the points on a *closed*, undirected line with respect to a tetradic relation $ABCD$ obtaining between points A, B, C, D in virtue of the separation of the pair BD by the pair AC . Axioms for separation of point pairs and for o-betweenness are given in E. V. Huntington: "Inter-Relations Among the Four Principal Types of Order," *Transactions of the American Mathematical Society*, Vol. XXXVIII (1935), pp. 1-9. Cf. also E. V. Huntington and K. E. Rosinger: "Postulates for Separation of Point-Pairs (Reversible Order on a Closed Line)," *Proceedings of the American Academy of Arts and Sciences* (Boston), Vol. LXVII (1932), pp. 61-145, and J. A. H. Shepperd: "Transitivities of Betweenness and Separation and the Definitions of Betweenness and Separation Groups," *Journal of the London Mathematical Society*, Vol. XXXI (1956), pp. 240-48.

tained by it but only the convention that a lower time number was assigned to it. By contrast, in the context of irreversible processes, the following results obtain: first, the relation term "earlier than" names an *objective* physical relation between two states which is *different* from the converse objective relation named by "later than," and second, the designation of *one* of two causally connected events as "the cause" of the other names a physically different relational attribute from the one named by the converse designation "the effect." Let us return to our reversible genidentical causal process of the rolling ball from whose description we have eliminated all reliance on the anisotropy of time. It is now clear that this renunciation of reference to all attributes depending on the anisotropy of time renders the rudimentary *causal* relation uniting the events of our genidentical reversible causal chain *symmetric*.⁸ And the properties of that *symmetric* causal relation would exhaust all the properties of causality in a world all of whose processes are de facto reversible.

Is it possible to provide an explicit definition of our symmetric causal relation without using any of the temporal concepts which that relation is first intended to define? Every attempt to do so known to or made by the writer has encountered insurmountable difficulties which are closely related to those familiar from the study of lawlike subjunctive conditionals by N. Goodman and others.⁹ Suppose, for example, that we attempted to define the symmetric causal relation between two genidentically related events E and E' by asserting that either of these two events is existentially a sufficient and a necessary condition for the other's occurrence in the following sense: if the set U of events constituting the universe contains an event of the kind E, then it also contains an event of the kind E', and if U does not contain an

⁸ In Chapter Eight, we shall give reasons for rejecting the objections to this conclusion set forth by Reichenbach in *DT*, *op. cit.*, p. 32.

⁹ Cf. N. Goodman: *Fact, Fiction and Forecast* (Cambridge: Harvard University Press; 1955), esp. pp. 13-31.

It is noteworthy that even in a context which, unlike ours, does presuppose the temporal concept of "later," C. I. Lewis reaches the conclusion (cf. his *Analysis of Knowledge and Valuation* [LaSalle: Open Court Publishing Co.; 1946], pp. 226-27) that the "if...then" encountered in causal statements does not yield to analysis, and therefore he speaks of the undefined "if...then of real connection."

E-like event, then it also will not contain an E'-like occurrence, no assumptions being made at all as to one of these two events being *earlier* than the other in the sense of presupposing the anisotropy of time. But this attempt at definition won't do for several reasons as follows: (1) the statement of the existential sufficiency of E is either tautological or self-contradictory, depending upon whether it is understood in its antecedent that the set U does or does not contain E', (2) the corresponding statement of existential necessity is either self-contradictory or tautological depending on whether in its antecedent, U is or is not construed as having E' among its members, and (3) attempting to turn these tautological or self-contradictory statements into true synthetic ones by making these assertions not about U itself but about appropriate proper *subsets* of U founders on (a) the need to utilize *temporal* criteria to circumscribe the membership of these subsets and (b) our inability to specify *all* of the *relevant conditions* which must be included in the antecedent, if the statement of the existential sufficiency of E is to be true.

The numerous difficulties besetting the specification of the relevant conditions¹ are not removed but only baptized by giving them a name—to borrow a locution from Poincaré—by the physicist's reference to the total state of a *closed system* in the antecedents of his causal descriptions of physical processes. Instead of enlightening us concerning the content of the *ceteris paribus* assumption, the invocation of the concept of closed system merely shifts the problem over to ascertaining the *cetera* which must materialize throughout the vast Minkowski light cones *outside* the system or on the spatial enclosure of the system in order that the system be *closed*.

These considerations suggest that we introduce the symmetric causal relation under discussion as a *primitive* relation for the purpose of then defining temporal betweenness and simultaneity. The reader will ask at once why the reduction of these ordinal concepts of time to such a primitive is not to be rejected as demanding too high a price epistemologically. Several weighty replies can be given to this question as follows:

(1) Despite the difficulties besetting the various versions of

¹ Many of these difficulties are discussed searchingly by N. Goodman: *Fact, Fiction and Forecast*, *op. cit.*, pp. 17-24.

the causal theory of time beginning with Leibniz's, the Einstein-Minkowski formulation of special relativity has sought to ground the temporal order of the physical world on its causal order by its reliance on influence chains (signals). Although the theory of relativity, like any geometry or other scientific theory, can be axiomatized in several different ways (at least in principle) by the use of different sets of primitives, the axiomatization of the relativistic topology of time on the basis of causal (signal) chains gives telling testimony of the desirability of formulating a philosophically viable version of the causal theory of time.

(2) The very explanation of some of the temporal features of the physical world on the basis of its causal features and the embeddedness of man's organism in this causal physical world lead us to expect that the causal acts of intervention (even insofar as they are reversible) which enter into man's *testing* of the Einstein-Minkowski theory will be part of the temporal order and will require for their practical execution by us *conscious* organisms recourse to the deliverances of our *psychological* sense of temporal order.

(3) My motivation for advancing below a particular version of the causal theory of time in which the attempt is made to dispense entirely with these psychological deliverances in the axiomatic foundations derives from the following two premises: (a) the thesis of astrophysics (cosmogony) and of the biological theory of human evolution that temporality is a significant feature of the physical world independently of the presence of man's conscious organism and hence might well be explainable as a purely physical attribute of those preponderant regions of space-time which are not inhabited by conscious organisms, and (b) the view of philosophical naturalism that man is part of nature and that those features of his conscious awareness which are held to be isomorphic with or likewise ascribable to the inanimate physical world must therefore be explained by the laws and attributes possessed by that world independently of human consciousness. When coupled with certain results of statistical thermodynamics, the version of the causal theory of temporal order to be advanced below provides a unified account of certain basic features of physical and psychological time. Since man's body participates in those purely physical processes which con-

fer temporality on the inanimate sector of the world and which are elucidated in part by the causal theory of time, that theory contributes to our understanding of some of the traits of psychological time.

So much for the justification of our use of the rudimentary symmetric causal relation as a primitive and of eschewing reliance on the deliverances of psychological time.²

Now if we are confronted with a situation in which two actual events E and E' are *genidentical* and hence causally connected in our rudimentary *symmetric* sense—or “k-connected” as we shall say for brevity—then we are able to use the properties of genidentical chains, which include causal continuity, to define both temporal betweenness and simultaneity.

For reasons which will become apparent, the causal definition of temporal betweenness to be offered will be given, pending the introduction of further requirements, so as to allow that time be topologically either open in the sense of being a system of o-betweenness like a Euclidean straight line, or closed in the sense of being a system of separation closure like a circle. In order to make the statement of the definition correspondingly general, we require the following preliminaries in which we use the abbreviation “iff” for “if and only if”:

(1) We shall call the quadruplet of events E L E' M (where

² The renunciation of essential recourse to *phenomenalistic* relations in the conceptual elaboration of the theory of temporal order was championed by Reichenbach in 1928, when he wrote (*Philosophie der Raum-Zeit-Lehre*, *op. cit.*, p. 161; cf. also pp. 327–28): “it is in principle impossible to use subjective feelings for the determination of the [temporal] order of external events. We must therefore establish a different criterion.” But in his last publication (*DT*, *op. cit.*, pp. 33–35) he invokes *direct observation* of nearby quasi-coincidences as a basis for giving meaning to the local order of temporal betweenness of such coincidences. If this observational criterion is intended to involve man’s subjective time sense, then one wonders what considerations persuaded Reichenbach to abandon his earlier opposition to it. And it is not clear how he would justify having recourse to it in the very context in which he claims to be providing a “causal definition” of the order of temporal betweenness “by means of reversible processes” (*DT*, *op. cit.*, p. 32). However, if he is making reference here not to the temporal deliverances of consciousness but rather to the directly observable indications of a material clock, then he has merely displaced the difficulty by posing but not solving the problem of showing how the causal features of a reversible clock furnish the definition of temporal betweenness.

$E \neq E'$) an "n-quadruplet," iff given the actual occurrence of E and E' , it is necessary that either L or M occur in order that E and E' be k -connected, L and M being genidentical with E and E' , and "or" being used in the inclusive sense. It is essential to note that "it is necessary that either L or M occur" does not entail "either the occurrence of L is necessary or the occurrence of M is necessary." The sentence " E L E' M are an n-quadruplet" will be abbreviated to " $n(E L E' M)$."

(2) We utilize the property of causal continuity possessed by genidentical causal chains and assert: for any two genidentically related (and hence k -connected) distinct events E and E' , there exist sets $\hat{X}$ and $\hat{F}$ of events genidentical with them, each of which has the cardinality $\aleph$ of the continuum and such that for each X belonging to $\hat{X}$ and each F belonging to $\hat{F}$, we have $n(E X E' F)$. Thus, there are $\aleph$ n-quadruplets.

(3) Given E and E' , we shall call a set α an "n-chain" connecting E and E' , iff the members X of α are given by the following condition:

$$X \in \alpha \text{ iff } (\exists F) [n(E X E' F) \cdot \sim n(E F E' F)].$$

All the members of n-chains connecting a given pair of genidentical events E and E' are thus genidentical with E and E' .

See Append. § 13

We are now able to define *temporal betweenness* as follows: any event belonging to an n-chain connecting a pair of genidentical events E and E' is said to be temporally between E and E' .

³ It might be objected that there is a difficulty in applying this definition to our earlier paradigmatic example of the rolling ball, for it would be possible to have someone place the ball at point P at the appropriate time even though that ball never was and never will be at the point P and to have a different ball placed appropriately at the latter point. Or a critic might say that even if the original ball were to be at one of the points P or P' , it could always be prevented from reaching the other point by being suitably intercepted while in transit. The aim of such objections would be to show that in order to rule out these alleged counter-examples, our construction would become circular by having to invoke temporal concepts which it is avowedly not entitled to presuppose. But the irrelevance of these objections becomes clear, when cognizance is taken of the fact that instead of exhibiting a circularity in our causal definition of temporal betweenness, they merely call attention to the existence of pairs of events which do not fulfill the conditions for the applicability of our definition, because they are not causally connected (or connectible) in the requisite genidentical way.

Once a definition of topological simultaneity becomes available, the definition just given admits of a generalization so as to allow that events which are not genidentical with E and E' but which are simultaneous with any event both temporally between and genidentical with them will likewise be temporally between E and E'. These topological definitions can then be particularized by the metrical definitions of simultaneity used in particular reference frames.

A very important feature of our definition of temporal betweenness is that it leaves open the question as to which one of the following alternatives prevails:

(1) The n-quadruplets which E and E' form with pairs of members of the n-chains connecting E and E' have the formal properties of the tetradic relation of separation, thus yielding a system of separation closure as the temporal order, or

(2) The membership of the n-chains connecting E and E' is such as to yield a system of o-betweenness as the temporal order.

In order to articulate this neutrality of the definition, we note, for purposes of comparison, the following *partial* sets of properties of the two alternative types of order in question. Letting "ABC" denote the triadic relation of betweenness, "ABCD" the tetradic relation of separation and "→" the relation of logical entailment, we have:

o-betweenness

1. $ABC \rightarrow CBA$ (symmetry in the end-points or "undirectedness")
2. $ABC \rightarrow \sim BCA$ (preclusion of closure)

separation-closure

If all four elements A, B, C, and D are distinct, then

1. $ABCD \rightarrow DCBA$ ("undirectedness")
2. $ABCD \rightarrow BCDA$ ("closure")

The corresponding *partial* properties of *cyclic* betweenness, as exemplified by the class of points on a *directed* closed line, are:

cyclic betweenness

1. $ABC \rightarrow \sim CBA$ (preclusion of symmetry in the endpoints or "directedness")
2. $ABC \rightarrow BCA$ (closure)

But we are not concerned with the species of closed order represented by cyclic betweenness, since we shall argue that it is not relevant to temporal betweenness.

The neutrality which we have claimed for our definition of temporal betweenness in regard to both o-betweenness and separation closure will be clarified by dealing with an objection which Dr. Abner Shimony has suggested to me for consideration.

If the definition is to allow time to be a system of o-betweenness and the subsequent introduction of a serial temporal order in which the members E, X, Y, E', F, and G of a genidentical causal chain are ordered as shown by the order of their names, then our definition ought not to entail that events like F, which are *outside* the time interval bounded by E and E', are temporally between E and E'. Nor should it entail that *inside* events like X and Y turn out not to be temporally between E and E'. Now, our definition of the membership of α chains was designed to preclude precisely such entailments as well as to allow the order of separation closure, in which every event is temporally between every other pair of events. The question is whether it could not be objected that we have failed nonetheless on the following grounds: events such as F and G, which are outside the interval EE', are each necessary for the respective occurrences of E and of E', just as much as inside events like X and Y are thus necessary; and does it not follow then that every "outside" event like F does enter into an n-quadruplet $n(E F E' F)$, thereby qualifying, no less than do "inside" events in this case, as necessary for the k-connectedness of E and E'?

If this conclusion did in fact follow, then our definition would indeed preclude o-betweenness by making the α -chains empty. For in that case, there would be no genidentical event whatever satisfying the requirement of not being necessary for the k-connectedness of E and E', as demanded by our definition of α . But the reasoning of the objection breaks down at the point of inferring that outside events like F and G do form n-quadruplets with E and E'. For although all events genidentical with E and E'—be they inside ones or outside ones—are necessary for the respective occurrences of E and of E', only the inside ones have the further property of being necessary for the k-connectedness of E and E', given that the latter do occur. By noting the distinc-

tion between the properties of (1) being necessary for the respective occurrences of E and E' and (2) being necessary for the k -connectedness of events E and E' whose occurrence is otherwise granted, we see that the objection to our definition derived its plausibility but also its lack of cogency from inferring the second property from the first.

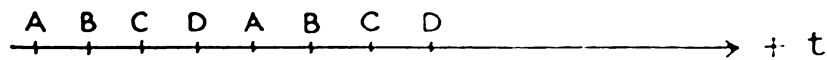
The fact that our definition of temporal betweenness does not itself discriminate between closed and open time becomes further evident upon considering universes to each of which it applies but which have topologically different kinds of time.

(A) CLOSED TIME

Let there be a universe consisting of a platform and one particle constantly moving in a circular path without friction. And be sure not to introduce surreptitiously into this universe either a conscious human observer or light enabling him to see the motion of the particle. Then the motion might be such that the temporal betweenness which it defines would exhibit closedness rather than openness, because there would be no physical difference whatever between a given passage of the particle through a fixed point A and its so-called "return" to the same state at A : instead of appearing periodically at the same place A at different instants in open time, the particle would be "returning"—in a highly Pickwickian sense of that term—to the selfsame event at the same instant in closed time. This conclusion rests on Leibniz's thesis that if two states of the world have precisely the same attributes, then we are not confronted by distinct states at different times but merely by two different names for the same state at one time. And it is this Leibnizian consideration which renders the following interpretation inadmissible as an alternative characterization of the time of our model universe: the same kind of set of events (circular motion) keeps on recurring eternally, and the time is topologically open and infinite in both directions. The latter interpretation is illegitimate since a difference in identity is assumed among events for which their attributes and relations provide no basis whatever. Hence that interpretation cannot qualify as a legitimate rival to our assertion that the events of our model world

form an array which is topologically closed both spatially and temporally.

Ordinary temporal language is infected with the assumption that time is open, and a description of the closed time of our model world in that language would take the misleading form of saying that the same sequence of states keeps recurring all the time. This description can generate pseudo-contradictions or puzzles in this context, because it suggests the following structure.



Here there are distinct sets of events ABCD which are merely the same in kind with respect to one or more of their properties. But this is not the structure of closed time. A closed time is very counterintuitive psychologically for reasons which will emerge later on. And hence the assumption of the closedness of time is a much stranger one intuitively than is a cyclic theory of history. In a cyclic theory of history, one envisions the periodic recurrence of the same kind of state at different times. And this conception of cyclic recurrence affirms the openness of time. Perhaps the lack of psychological imaginability as distinct from theoretical intelligibility of a closed time accounts for the fact that its logical possibility is usually overlooked in theological discussions of creation. This failure of imagination is unfortunate, however, since there could be no problem of a beginning or creation if time were to be cosmically closed.

Three kinds of objections might be raised to my claim that our model universe does provide a realization of a topologically closed kind of time: First, it might be argued that my earlier caveat concerning the need for a highly Pickwickian construal of the term "returning" actually begs the question in the following sense: the mere contemplation of the model universe under consideration compels the conclusion that the particle does indeed return in the proper sense of that term to the same point A at different instants of open time. And it might be charged that to think otherwise is to disregard a plain fact of our contemplative experience. Second, the complete circular motion could be subdivided into a finite number n of (equal)

parts or submotions (episodes). And this has been claimed to show that instead of being topologically closed, the time of our model universe has the open topology of a finite segment of a straight line which is bounded by distinct endpoints, the first and n^{th} submotions allegedly being the ones which contain the two terminal events. And third, it has been said that it is of the essence of time to be open. Noting that my characterization of the model universe at issue as having a closed time depends on the invocation of Leibniz's principle of the identity of indiscernibles, a critic has maintained that instead of showing that a closed time is logically possible, this model shows that Leibniz's principle must be false!

As to the first of these objections, which rests on the deliverances of our contemplation of the model universe, my reply is that the objector has tacitly altered the conditions that I had postulated for that model universe, thereby tampering with the very features on which my affirmation of the closure of its time had been predicated. For the objector has not only introduced a conscious organism such as himself, whom he presumes to have distinct memories of two passages at A, but he has also surreptitiously brought in another physical agency needed to make these distinct memories possible: a light source such as a candle which enables him to see and which distinguishes an earlier and a later passage at A by being more dissipated or burnt-out at the time of the later passage. Thus, the objection is untenable, because the objector assumes a universe differing from my hypothetical one so as not to have a closed time.

The second objection, which adduces the n submotions, is vitiated by the following gratuitous projection of the ordinal properties of numerical names onto the events (or submotions) to which they are assigned in a counting procedure: the divisibility of the motion of the particle on the platform into n subintervals of events—where n is a cardinal number—does not make for the possession of any objective property of being *temporal termini* by the particular subintervals which were assigned the numbers 1 and n respectively. For no two of the n subintervals of events—whatever the particular numbers that happened to have been assigned to them in the counting—are ordinally distinguished qua termini from any of the others. If therefore we

count them by arbitrarily assigning the number 1 to some one of them and by then assigning the remaining $n-1$ numbers, this cannot serve to establish that temporally the ordinal properties of the particular subintervals thus accidentally named 1 and n respectively are any different from those of the subintervals which are thereby assigned natural numbers between 1 and n .

The justification for this rebuttal can be thrown into still bolder relief by noting the objective differences between our model universe of the particle moving "perpetually" in a circular path and the following different model universe whose time does indeed have the open topology of a finite segment of a straight line bounded by two endpoints: a platform universe differing from the first one only by exchanging the particle moving in a closed (circular) path for a simple pendulum oscillating "perpetually" and frictionlessly over the platform as if under the action of terrestrial gravity. Let the oscillation be through a fixed small amplitude $\theta = 2\alpha$ between two fixed points whose angular separation from the vertical is $+\alpha$ and $-\alpha$ respectively. Then the finitude of the time of this pendulum universe is assured by the fact that the "perpetual periodic returns" of the pendulum bob to the same points over the platform are, in fact, identical events by Leibniz's principle. For this identity prevents this latter model universe from qualifying as an infinitely periodic universe whose time is open. But the two fixed points at angular distances $-\alpha$ and $+\alpha$ respectively from the vertical uniquely confine the motion spatially to the points between them. And hence the events constituted by the presence of the pendulum bob at $-\alpha$ and $+\alpha$ are objectively distinguished as termini from all other events belonging to the motion of the pendulum, although no one of these two events is distinguished from the other as the first rather than the last, since the motion is reversible. Accordingly, the time of the pendulum universe has the topology of an undirected finite segment of a straight line bounded by two termini. But there is no foundation for the objection that our first model universe of the particle in the circular path has the latter kind of time.

The third objection, which has the spirit of an argument based on Kant's presuppositional method, suffers from precisely the

same well-known logical defects as does the claim that we know a priori that the universe must be spatially infinite (topologically open) rather than finite (closed but unbounded). For the defender of the a priori assertibility of the infinite Euclidean topology would offer the following corresponding argument regarding space: if it appeared that a spatial geodesic of our universe were traversable in a finite number of equal steps terminating at the same spatial point, the certification of the sameness of that point via Leibniz's principle would have to be rejected. And the a priori proponent of the Euclidean as opposed to the spherical topology of space would then adduce these circumstances to show that Leibniz's principle is false. Although the rise of the non-Euclidean geometries has issued in the displacement of the Kantian conception of the topology of physical *space* by an empiricist one in most quarters, a vestigial Kantianism persists in many quarters with respect to the topology of *time*. And this lingering topological apriorism seems to be nurtured by the following failure of imagination: the neglect to envision having to divest the topology of the time of a model universe—or the cosmic time of our actual universe—of some of the topological properties of the cosmically local time of our everyday experience. Such divestiture would be necessary, if Einstein's field equations were to allow solutions yielding temporally closed geodesics, as has been claimed by Gödel.⁴ If correct, Gödel's result would seem to indicate that temporally closed geodesics can also be possessed by a deterministic universe containing sentient beings like ourselves. But Chandrasekhar and

Correction in §15 of
the Appendix

⁴ In recent papers ("An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation," *Reviews of Modern Physics*, Vol. XXI [1949], p. 447, and "A Remark About the Relationship of Relativity Theory and Idealistic Philosophy," in P. A. Schilpp [ed.] *Albert Einstein: Philosopher-Scientist*, *op. cit.*, pp. 560–62), Gödel has maintained that there exist solutions of Einstein's field equations which assert the existence of *closed* timelike world-lines. Einstein says (*ibid.*, p. 688) that "It will be interesting to weigh whether these [solutions] are not to be excluded on physical grounds." Reichenbach himself points out (*PST*, *op. cit.*, 141, 272–73) that in a world of closed timelike world lines, his causal criterion of temporal order becomes self-contradictory in the large (*Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, p. 22). And, as Einstein remarks (*ibid.*), in such a world, irreversibility also fails to hold in the large.

Wright⁵ have argued very recently that Gödel's claim is mathematically unfounded.

There is a rather simple way of seeing how manlike beings might discover that the cosmic time of their universe is closed, despite the seriality of the local segment of cosmic time accessible to their daily experience. Suppose that all the equations governing the temporal evolution of the states of physical and biological systems are deterministic with respect to the properties of events and that these equations are formulated in terms of a time variable t ranging over the real numbers, it thereby being assumed to begin with that time is topologically open. Now postulate further that the boundary conditions of this deterministic world are such that all of the variables of state pertaining to it (including those variables whose values characterize the thoughts of scientists living in it) assume precisely the same values at what are *prima facie* the time t and a very much later time $t + T$. Then upon discovering this result by calculations, these scientists would have to conclude that the two different values of the time variable for which this sameness of state obtains do not denote two objectively distinct states but are only two different numerical names for what is identically the same state. In this way, they would discover that their universe is temporally closed, much as a scientist who begins by assuming that the universe is spatially infinite may then find that, in the large, it is spatially closed. But there is an important difference between the psychological intuitability of a closed space and that of a closed time: a cosmically closed time could *not everywhere* be locally serial with respect to "earlier than."

As previously emphasized, it is of crucial importance in this context, if pseudo-puzzles and contradictions are to be avoided, that the term "returning" and all of the preempted temporal language which we tend to use in describing a world whose time (in the large) is closed be divested of all of its tacit reference to an external serial super-time. Awareness of the latter pitfall now enables us to see that as between the two kinds of closed order which we have mentioned, separation closure and not

⁵ S. Chandrasekhar and J. P. Wright: "The Geodesics in Gödel's Universe," *Proceedings of the National Academy of Sciences*, Vol. XLVIII (1961), pp. 341-47, and esp. p. 347.

cyclicity must be held to be the order of a closed time. For in the context of physical states, cyclic betweenness depends for its directional anisotropy on an appeal to a serial and hence open time and derives its closure only from a spatial periodicity. Thus, a closed physical time must exhibit separation-closure and cannot meaningfully be cyclic. And if we are to give a concise characterization of a physically plausible closed time, it should read: every state of the world is temporally between every other pair of states of the universe in a sense of "between" given by our definition for the case of separation closure. It might be asked why we have been assuming that the structure of a closed time would have to be that of a knot-free circle rather than that given by the self-intersecting closed line in the numeral 8. The reply is that the framework of our models of a closed time is deterministic and that the course of the phase curve representing a finite (closed) mechanical or other deterministic system is uniquely determined by any one of its phase points.

See Appendix §16

Our characterization of a closed time is to apply not only to a thoroughly uninteresting model world like that of the solitary particle moving in a circular path on a platform but to a cosmos whose total states are not elementary events but large classes of coincidences of many genidentical objects. It is therefore essential that we specify the meaning of topological simultaneity, which is presupposed by the concept of a state of the world. Reserving comment until we discuss open time and, in particular, the species of open time affirmed by the special theory of relativity, on the serviceability of the following definition in that context, we define: two events are *topologically simultaneous*, iff it is not physically possible that they be connected by genidentical causal chains. It is to be noted in connection with this definition that a light ray directly connecting a pair of macroscopic events is held to qualify as an entity possessing genidentity, although this assumption no longer holds in contexts in which the Bose-Einstein statistics applies to photons.

(B) OPEN TIME

Since our definition of topological simultaneity is completely neutral as to the closedness or openness of time, it is apparent

that we can also utilize the definition of topological simultaneity for the description of a universe whose time is characterized by o-betweenness, such as that of Newton or of special relativity. In the latter Einsteinian world, the limiting role played by electromagnetic causal chains makes for the fact that the topological definition of simultaneity leaves a good deal of latitude for the synchronization rule and thereby for the metrical definition of simultaneity. And hence in that case, our purely ordinal definition would have to be particularized in each Galilean frame, as we shall see in detail in Chapter Twelve, so as to render metrically simultaneous in any given frame only those pairs of causally non-connectible events which conform to the criterion of that particular frame.

To appreciate the role of boundary conditions in conferring openness on time, we first recall our pendulum world model of a finite open time and now proceed to provide models of an infinite open time as follows: first a very simple kind of world and then a world having far greater relevance to the actual world in which we live.

Let there be a universe consisting of a platform, material clocks and at least two simple pendulums X and Y which have incommensurable periods of oscillation so that after once being in the same phase they are permanently out-of-phase with each other. Then these motions would define an infinite open time in virtue of Leibniz's non-identity of discernibles: in this case, a given passage E_p of the bob of pendulum X through a fixed point P would be physically different from any other such passage E_i in virtue of E_p 's being simultaneous with a different phase of pendulum Y from the one with which E_i is simultaneous, thereby giving rise to an order of o-betweenness for time. In the over-all light of the construction of physical time presented in this book, this assertion involves a philosophical commitment to a Leibnizian criterion for the individuality of the events belonging to the causal chains formed by genidentical classical material particles or macro-objects, along with a non-Leibnizian primitive concept of material genidentity for the entities whose relations generate these events. I do not see any inconsistency or circularity in this feature of the construction. In particular, it seems to me that only a confusion of the context of justification with

the context of discovery (in Reichenbachian terminology) or of the factual reference with the evidential base (in Feigl's parlance) can inspire the charge that it is circular to use the concept of (material) genidentity as a primitive in a reconstruction of the temporal order of physical events, on the alleged grounds that the meaning of the temporal order is already presupposed in our *recognitions* of objects as the same upon encountering them in different places at different times.

Now consider a finite universe or a large finite quasi-closed portion of our actual universe, if the latter be spatially infinite, to which the Maxwell-Boltzmann gas statistics is roughly applicable. The concept of what constitutes an individual micro-state ("arrangement" or "complexion") in the Maxwell-Boltzmann statistics depends crucially on the assumption that material genidentity can be ascribed to particles (molecules) and involves the very Leibnizian conception of the individuality of events which we set forth by reference to our illustrative examples of the simple model universes. An individual instant of time is thus defined for this universe by one of its particular micro-states. And, on this criterion, it will therefore be quite meaningful to speak, as I shall later on, of the occurrence of the same *macro*-state ("distribution") and hence of the same entropy at different times, provided that the respective underlying *micro*-states are different. But whether or not a universe constituted by a Maxwell-Boltzmann gas will exhibit a set of micro-states which define an open time rather than a time which, in the large, is closed depends not on the causal character of the motions of the constituent molecules but on the boundary conditions governing their motions! And whether the time thus defined will, if open, also be infinite, depends on the micro-states having the degree of specificity represented by points in phase space rather than by the mere cells used to compute the coarse-grained probabilities of various macro-states. For a finite closed mechanical system of constant energy is at least quasi-periodic and can possibly qualify as aperiodic only with respect to a punctual characterization of its micro-states.

Hence, the symmetric kind of causality affirmed by the equations of mechanics themselves, as distinct from prevailing nomologically contingent boundary conditions to which they apply,

allows but does not assure that the temporal betweenness (and simultaneity) defined by genidentical causal chains is that of an open rather than a closed ordering.⁶

This analysis of the physical basis of open time requires the addition of a comment on the meaning of the "reversibility" of mechanical motions. We observed that the total states of our pendulums *X* and *Y* whose periods are incommensurable define an infinite open time by not exhibiting any "reversals." What is meant, therefore, when we attribute reversibility to the motions of either of the individual pendulums is that either elementary constituent of the total system can (under suitable boundary conditions of its own) give rise to the same kind of event at different times. We do not mean that the pendulum in question has "returned" to the selfsame event, since the total states of the complete physical system assure via the Leibnizian non-identity of discernibles that the events belonging to the individual pendulum form an infinite open order of time. Thus, the reversibility of the laws of mechanics has a clear meaning in the context of an infinite open time. And the reversibility of the elementary processes in our Maxwell-Boltzmann universe, which is affirmed by these laws, is therefore entirely compatible with the infinite openness of the time defined by the total micro-states of that universe. We shall soon see, however, that the mere *non-reversal* of the *total* micro-states, which assures the infinite openness of time, is a much weaker property of these states than the kind of *irreversibility* that is a sufficient condition for the anisotropy of time.

We have emphasized the neutrality of our definition of temporal betweenness on the basis of causal continuity in regard to

⁶ Mehlberg ["Essai sur la théorie causale du temps," *op. cit.*, Vol. I, p. 240] correctly calls attention to the fact that the principle of causal continuity is independent of whether physical processes are reversible or not. But then, affirming the complete reversibility of the physical world and thereby the isotropy of physical time (*ibid.*, p. 184), he takes it for granted that the principle of causal continuity as such always defines an open betweenness (*ibid.*, pp. 239-40, and *op. cit.*, Vol. II, pp. 179, 156-57, 168-69). But, as we just saw, in a completely reversible world, it can happen that the betweenness defined by the principle of causal continuity for genidentical chains is that of the closed variety associated with separation closure. Thus, instead of being isotropic while being open, time would then be both isotropic and closed.

the rival possibilities of open and closed betweenness for time. The second feature of our definition of temporal betweenness which needs to be noted is that it utilizes the concept of physical possibility and thereby runs the logical risk mentioned earlier in this chapter, *if* we expect it to order B as temporally between A and C even in those cases where B would be between A and C if A and C *were* genidentically connected but are actually not. This risk also applies to our definition of (topological) simultaneity for noncoincident events.

In the theories of pair-production due to Wheeler, Feynman, and Stückelberg,⁷ some of the phenomena under investigation may be described in the usual macro-language by saying that a "particle" can travel both "forward" and "backward" in macro-time. Thus, in that description a "particle" can violate the necessary condition for simultaneity given in our definition here by being at two different places at instants which are macroscopically simultaneous. But this fact does not disqualify our definition of simultaneity. For the topology of the time whose physical bases our analysis is designed to uncover is defined by (statistical) *macro*-properties for which these difficulties do not arise. The macro-character of our concept of simultaneity is evident from the fact that it depends on the concept of material or quasi-material genidentity. Precisely this concept and the associated classical concept of a particle trajectory are generally no longer applicable to micro-entities, so that the consequences of their inapplicability in the Wheeler-Feynman theory and in the Bose-Einstein statistics need not occasion any surprise. Furthermore, the macro-character of the anisotropy of time will emerge from the analysis to be given later in Chapter Eight.

We see that since our definitions of temporal betweenness and simultaneity employed a concept of causal connection joining two events which made no reference to *one* of the events being *the* cause of the other by being *earlier*, these definitions do not presuppose in any way the anisotropy of time. Nevertheless, they do exhaust the contribution which the causality of reversible processes can make to the elucidation of the structure of time. If physical time is to be *anisotropic*, then we must look to fea-

⁷ Cf. H. Reichenbach: *DT*, *op. cit.*, pp. 262-69.

tures of the physical world other than the causality of irreversible processes as the source. In particular, since it has been shown that if the micro-statistical analogue of entropy fails to confer anisotropy on time, all other micro-statistical properties of closed systems for which an entropy is defined will fail as well, we must turn to an examination of entropy to see in what sense, if any, it can supply attributes of physical time not furnished by causality.

Before doing so, however, I wish to make a concluding comment on the causal theory of time in its bearing on the modern mathematical resolution of Zeno's paradoxes of motion.

Having no recourse to the anisotropy or even to the openness of time, our definitions of temporal betweenness and simultaneity established a *dense* temporal order. For our construction entails that between any two events, there is always another. Now, as I have explained elsewhere in 1950,⁸ it was the ascription of this denseness property to the temporal order by the modern mathematical theory of motion which prompted the Zenonian charge by such philosophers as William James and A. N. Whitehead that this theory of motion is neither physically meaningful nor consistent. Resting their case on the immediate deliverances of consciousness, which include "becoming," these philosophers maintained that the temporal order is discrete rather than dense, the events of nature occurring *seriatim* or, as Paul Weiss put it, "pulsationally."¹ We see that the causal theory of time as here presented refutes their polemic on the issue of the denseness of the temporal order of the physical world and that this refutation is not dependent on the logical viability of Reichenbach's unsatisfactory mark method, which I employed in my paper of 1950, when giving a critique of their arguments on the basis of the causal theory of time.²

⁸ Cf. A. S. Eddington: *The Nature of the Physical World* (New York: The Macmillan Company; 1928), pp. 79-80. For details on the "principle of detailed balancing" relevant here, cf. R. C. Tolman: *The Principles of Statistical Mechanics* (Oxford: Oxford University Press; 1938), pp. 165, 521.

⁹ A. Grünbaum: "Relativity and the Atomicity of Becoming," *The Review of Metaphysics*, Vol. IV (1950), pp. 143-86.

¹ Cf. Paul Weiss: *Reality* (New York: Peter Smith; 1949), p. 228.

² The analysis on the basis of the mark method is given on pp. 160-86 of my previously cited "Relativity and the Atomicity of Becoming." For a treatment not encumbered by the weaknesses of the mark method, cf. A. Grünbaum: "Modern Science and the Refutation of the Paradoxes of Zeno," *The Scientific Monthly*, Vol. LXXXI (1955), pp. 234-39.

Chapter 8

THE ANISOTROPY OF TIME

See
Append. §§11 and 20

(A) IS THERE A THERMODYNAMIC BASIS FOR THE ANISOTROPY OF TIME?

Just as we can coordinatize one of the dimensions of space by means of real numbers *without* being committed to the anisotropy of that spatial dimension, so also we can coordinatize a topologically open time-continuum without prejudice to whether there exist irreversible processes which render that continuum anisotropic. For so long as the states of the world (as defined by some one simultaneity criterion) are ordered by a relation of temporal betweenness having the same formal properties as the spatial betweenness on a Euclidean straight line, there will be two senses which are opposite to each other. And we can then assign increasing real number coordinates in one of these senses and decreasing ones in the other by convention *without* assuming that these two senses are *further distinguished* by the *structural* property that some kinds of sequences of states encountered along one of them are never encountered along the other.

If the latter situation does indeed obtain because of the existence of irreversible kinds of processes, then the time continuum is anisotropic. By the same token, if the temporal inverses of all kinds of processes actually materialized, then time would be isotropic.

We shall have to determine what specific properties of the physical world, if any, confer anisotropy on the time of nature in the sense of structurally distinguishing the opposite directions of "earlier" and "later" on the time axis in a specifiable way. We

See Append. §28

shall then be ready to see in Chapter Ten whether any feature of the universe such as the hypothetical one of indeterminism can give a physical meaning to the following further purported property of time: "passage," "coming into being," "becoming," "flux," or "flow," conceived as the *shifting* or transience of the "Now" along the structurally distinguished direction of "later than" on the time axis.

It is clear that the anisotropy of time resulting from the existence of irreversible processes consists in the mere structural differences between the two opposite senses of time but provides no basis at all for singling out *one* of the two opposite senses as "*the* direction" of time. Hence the assertion that irreversible processes render time anisotropic is *not at all* equivalent to such statements as "time flows one way." And the metaphor of time's "arrow," which Eddington intended to refer to the anisotropy of time, can be misleading: attention to the head of the arrow to the exclusion of the tail may suggest that there is a "flow" in *one* of the two anisotropically-related senses.

It will be essential to begin by giving an analysis of the concept of a temporally irreversible process in its bearing on the anisotropy of time with a view to then determining the following: to what extent, if any, do the kinds of physical processes obtain in our universe confer anisotropy on time?

In what sense do the failure of the dead to rise from their graves and the failure of cigarettes to reconstitute themselves from their ashes render the processes of dying and cigarette burning "irreversible"? There is both a weak sense and a strong sense in which a process might be claimed to be "irreversible." The weak sense is that the temporal inverse of the process in fact never (or hardly ever) occurs with increasing time for the following reason: certain particular de facto conditions ("initial" or "boundary" conditions) obtaining in the universe independently of any law (or laws) combine with a relevant law (or laws) to render the temporal inverse de facto non-existent, although no law or combination of laws itself disallows that inverse process. The strong sense of "irreversible" is that the temporal inverse is impossible in virtue of being ruled out by a law alone or by a combination of laws. In contexts calling for the distinction between these two senses of "irreversible," we shall use essentially

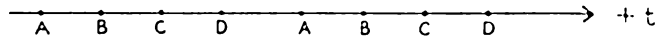
the terminology of H. Mehlberg,¹ and shall speak of the stronger, law-based kind of irreversibility as "nomological" while referring to the weaker kind of irreversibility as being "de facto" or "nomologically contingent." In the absence of these qualifications, the ascription of irreversibility to a process commits us to no more than the non-occurrence or virtual non-occurrence of its temporal inverse and leaves it open whether the irreversibility is de facto or nomological in origin. The neutrality of our use of the term "irreversible" as between the nomological and de facto senses is an asset in our concern with the anisotropy of time. For what is decisive for the obtaining of that anisotropy is not whether the non-existence of the temporal inverses of certain processes is merely de facto rather than nomological; instead, what matters here is whether the temporal inverses of these processes always (or nearly always) fail to occur, whatever the reason for that failure! Thus, the processes of masticating food and mixing cream with coffee are irreversible in this *neutral* sense. Hence if a silent film of a dinner party were to show a whole beef steak being reassembled from "desalivated" chewed pieces or the mixture of coffee and cream unmixing, we would know that it has been played backward.

The distinction between the weak and strong kinds of irreversibility has a clear relevance to those physical theories which allow a sharp distinction between laws and boundary conditions in virtue of the repeatability of specified kinds of events at different places and times. But it is highly doubtful that this distinction can be maintained throughout *cosmology*. For what criterion is there for presuming a spatially ubiquitous and temporally permanent feature of the universe to have the character of a boundary condition rather than that of a law?

Let us consider graphically the significance of the presumed fact that there are kinds of sequences of states ABCD, occurring with increasing time, such that the opposite sequence DCBA would not also occur with increasing time. Suppose, for example, ABCD in the diagram are successive kinds of states of a house that burns down completely with increasing or later time. Then

¹ H. Mehlberg: "Physical Laws and Time's Arrow," in: H. Feigl and G. Maxwell (eds.), *Current Issues in the Philosophy of Science*, *op. cit.*, pp. 105-38.

there will be no case of the inverse kind of sequence DCBA with increasing time, since the latter would constitute the resurrection of a house from debris. Thus, this opposite kind of sequence DCBA would exist only in the direction of decreasing or earlier time, while the first kind of sequence ABCD would not obtain in the latter direction.



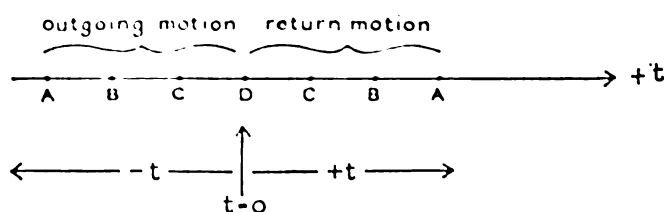
Accordingly, comparison of the structures of the opposite directions of time shows that, at least for the segment of cosmic time constituting the current epoch in our spatial region of the world, the kinds of sequences of states exhibited by the one direction are different from those found in the other. Hence we say that, at least locally, time is anisotropic rather than isotropic. It will be noted that the anisotropy of physical time consists in the mere structural differences between the opposite directions of physical time and constitutes no basis at all for singling out one of the two opposite directions as "the" direction of time.

The dependence of such anisotropy as is exhibited by time on the irreversible character of the processes obtaining in the universe can be thrown into still bolder relief by noting what kind of time there would be, if there were no irreversible processes at all but only reversible processes whose reversibility is not merely nomological but also *de facto*. That is to say the temporal inverses would not only be allowed by the relevant laws but would actually exist in virtue of the obtaining of the required initial (boundary) conditions. To forestall misunderstandings of such a hypothetical eventuality, it must be pointed out at once that our very existence as human beings having memories would then be impossible, as will become clear later on. Hence, it would be entirely misconceived to engage in the inherently doomed attempt to imagine the posited eventuality within the framework of our actual memory-charged experiences, and then to be dismayed by the failure of such an attempt. As well try to imagine the visual color of radiation in the infra-red or ultra-violet parts of the spectrum.

To see that if all kinds of natural processes were actually *de facto* reversible, time would indeed be *isotropic*, we now consider an

example of such a reversible physical process: the frictionless rolling of a ball over a path AD in accord with Newton's laws, say from A to D. This motion is both nomologically and de facto reversible, because (1) Newton's laws likewise allow another motion over the same path but from D to A, which is the temporal inverse of the motion from A to D, and (2) there are actual occurrences of this inverse motion, since the initial conditions requisite to the occurrence of the inverse motion do obtain.

Let us plot on a time axis the special case in which a particular ball rolls from A to D and is reflected so as to roll back to A, the zero of time being chosen for the event of the ball's being at the point D. The letters "A," "B," "C," and "D" on the time axis in our diagram denote the respective *events* of being at the point A in space, etc., thereby representing the sequence of states (events) ABCD of the "outgoing" motion and then the states DCBA of the "return" motion.



Mathematically, the nomological reversibility of the processes allowed by Newton's laws expresses itself by the fact that the form of the Newtonian equations of motion remains unaltered or invariant upon substituting $-t$ for $+t$ in them. We say, there-

² In the present context of exclusively reversible processes, the relation of "earlier than" implicitly invoked in the assertion that a ball moves "*from A to D*" (or, in the opposite direction, "*from D to A*") must be *divested* of its customary anchorage in an anisotropic time. For in the world of exclusively reversible processes now under discussion, the assertion that a given motion of a ball was "*from A to D*" rather than "*from D to A*" expresses not an objective physical relation between the two terminal events of the motion, but only the convention that we have assigned a lower time-number to the event of the ball's being at A than to its being at D. And this absence here of an objective physical basis for saying that the motion was "from" one of two points "to" the other makes for the fact, noted in Chapter Seven above, that, if all the processes of nature are de facto reversible, there is no objective physical basis for singling out one of two motion states of a ball as "the" cause of the other.

fore, that Newton's laws for frictionless motions are "time-symmetric." And hence our diagram shows that for every state of the ball allowed by Newton's laws at a time $+t$, these laws allow precisely the same kind of state at the corresponding time $-t$. In other words, in the case of reversible processes, the sequences of (allowed) states along the opposite directions of the time-axis are, as it were, mirror-images of each other. Hence, if all of the processes of nature were de facto reversible, time would be isotropic.

Thus, it is further apparent that the structure of time is not something which is apart from the particular kinds of processes obtaining in the universe. Instead, the nature of time is rooted in the very character of these processes.

Our remark in the last footnote on the logical status of the relation "earlier than" in a world of exclusively reversible processes must now be amplified by reference to Reichenbach's unsuccessful account of the logical difference between (1) that relation and the corresponding relation obtaining in a world containing irreversible processes, and (2) the corresponding two causal relations.

To begin with, it must be noted that while it is readily possible to define a triadic relation having all of the formal properties of o-betweenness in terms of a particular dyadic serial relation, the converse deduction is not possible, since in a given system of serial order, we can distinguish one "direction" from its opposite, whereas the system of o-betweenness does not, by itself, enable us to make such a differentiation. The case of the Euclidean straight line will illustrate this fact. The points of the straight line form a system of o-betweenness. This order is intrinsic to the straight line in the sense that its specification involves no essential reference to an external viewer and his particular perspective. The serial ordering of the points with respect to a concrete relation "to the left of" is extrinsic in the sense of requiring reference to an external viewer, at least for the establishment of an asymmetric dyadic relation "to the left of" between two given arbitrarily selected reference points U and V. Once we thus introduce an asymmetric dyadic relation between two such points, then, to be sure, we can indeed use the intrinsic system

of *o*-betweenness on the line to define a serial order throughout the line. To say that a given serial order on the line with respect to the relation "to the left of" is conventional is another way of saying that it is extrinsic in our sense. For a particular external perspective, it is of course not arbitrary whether a given point *x* is to the left of another point *y* or conversely. In contrast to the "extrinsic" character of the serial ordering of the points on the line with respect to the relation "to the left of," the serial ordering of the real numbers with respect to "smaller than" is intrinsic in our sense, since for any two real numbers, the ordering with respect to magnitude requires no reference to entities outside the domain.

It is essential not to overlook, as Reichenbach did and as the writer did in an earlier publication,⁴ that a serial ordering always establishes a difference in direction independently of whether it is intrinsic or extrinsic! Confusing extrinsicity of a serial relation with undirectedness, Reichenbach maintains that the relation "to the left of" on the line, though serial, is not "unidirectional," thereby allegedly failing to distinguish two opposite directions from one another, whereas he regards the relation "smaller than" among real numbers as both serial and "unidirectional."⁵ But he found himself driven to this contention only because he failed to note that, by being asymmetric, a serial relation is automatically a directed one even when the seriality has an extrinsic basis. And this oversight led him to distinguish relations which are serial while allegedly being undirected from directed serial relations. The latter error, in turn, issued in his

³ The relevant details on these formal matters can be found in Lewis and Langford: *Symbolic Logic* (New York: The Century Company; 1932), pp. 381–87, and in E. V. Huntington: "Inter-Relations Among the Four Principal Types of Order," *Transactions of the American Mathematical Society*, Vol. XXXVIII (1935), Sec. 3.1, p. 7.

⁴ H. Reichenbach: "The Philosophical Significance of the Theory of Relativity," in P. A. Schilpp (ed.) *Albert Einstein: Philosopher-Scientist*, *op. cit.*, pp. 304–05, and *DT*, *op. cit.*, pp. 26–27; A. Grünbaum: "Time and Entropy," *American Scientist*, Vol. XLIII (1955), p. 551.

⁵ For a telling critique of Reichenbach's account of the alleged general logical features of "unidirectional" as opposed to "merely" serial relations, cf. H. Mehlberg: "Physical Laws and Time's Arrow," in H. Feigl and G. Maxwell (eds.) *Current Issues in the Philosophy of Science*, *op. cit.*, pp. 109–11.

false distinction between the supposed mere seriality of time, which he called "order," and its "direction," a distinction which he attempted to buttress by pointing to the seriality of the time of Newton's mechanics and of special relativity in the face of the complete time-symmetry of the fundamental laws of these theories. But Reichenbach's distinction should be replaced by the distinction between intrinsically isotropic and anisotropic kinds of time, which we shall now explain.

De facto reversible processes intrinsically define a temporal order of mere o-betweenness under suitable boundary conditions, but the symmetric causal relation associated with these processes provides no physical basis for an intrinsic serial order of time. Just as it was possible, however, in the case of the Euclidean line, to introduce a serial ordering in its system of o-betweenness by means of an extrinsically grounded asymmetric dyadic relation between two chosen reference points U and V, so also it is possible to choose two reference states in a time that is intrinsically merely a system of o-betweenness, and extrinsically render this time serial by making one of these two states later than the other through the assignment of suitable real numbers as temporal names. It is in this sense that a world containing no kinds of irreversible processes can nonetheless be legitimately and significantly described by a serial time.

We see that if a universe contains no nomologically or de facto irreversible kinds of processes but is nonetheless claimed to have a serial time, this seriality must have an *extrinsic* component. For, in the presence of suitable boundary conditions, this hypothetically reversible universe will intrinsically define only a temporal order of o-betweenness. And this latter order is isotropic in the following twofold sense: first, all elementary processes are de facto reversible, and second, there is no one property of physical systems such as the entropy whose values intrinsically define a dyadic relation between pairs of states of these systems such that the class of states forms a serial order with respect to that relation.

But in a non-equilibrium world to which the non-statistical second law of classical thermodynamics is actually applicable, precisely the latter kind of property does exist in the form of the entropy. And hence such a world is temporally anisotropic: its

time exhibits a special kind of *difference in direction* arising from the directed, intrinsically grounded serial relation of "later than." It is apparent that if we say that the processes in such a world are "irreversible," our assertion differs logically from affirming the "non return" of any of the total states of a universe whose time is therefore open and perhaps also infinite (in one or both directions). For the *latter* kind of time can be *intrinsically isotropic*, and indeed will be isotropic if the universe possessing it contains only de facto reversible kinds of processes. By contrast, in the former, irreversible kind of universe, whose time is anisotropic, the following two features are present: first, the classical entropy law precludes the occurrence of the same (*non-equilibrium*) *macro-state* at different times rather than merely asserting that the *micro-states* define an *open* order of time in virtue of obtaining boundary conditions, and second, that law makes a specific assertion about the way in which macro-states occurring at different times do differ with respect to a single property.

Although the serial relation "later than" itself does have a "direction" in the obvious sense of being asymmetric, the set of states ordered by it does not have a direction but rather exhibits a special *difference* or anisotropy as between the *two opposite directions*. Thus, when we speak of the anisotropy of time, this must not be construed as equivalent to making assertions about "the" direction of time. J. J. C. Smart and Max Black have correctly pointed out⁶ that reference to "the" direction of time is inspired by the notion that time "flows." In particular, as we shall see in Chapter Ten, Reichenbach's assertions about "the" direction of time rest on his incorrect supposition that there is a physical basis for becoming in the sense of the shifting of a physically defined "now" along one of the two physically distinguished directions of time. By contrast, our characterization of physical time as anisotropic involves no reference whatever to a transient division of time into the past and the future by a "now" whose purported unidirectional "advance" would define "the" direction of time. In fact, we shall argue in Chapter Ten

⁶ J. J. C. Smart: "The Temporal Asymmetry of the World," *Analysis*, Vol. XIV (1954), p. 81, and M. Black: "The 'Direction' of Time," *Analysis*, Vol. XIX (1959), p. 54.

that the concept of becoming has no significant application to physical time despite its relevance to psychological and common sense time, because the "now" with respect to which the distinction between the past and the future acquires meaning depends crucially on the egocentric perspectives of a conscious organism for its very existence. Nevertheless, having entered this explicit caveat, we shall achieve brevity by using the locution "the direction of time" as a synonym not only for "the future direction" in psychological time but also for "the one of two physically distinguished directions of time which our theory calls 'positive.'"

Our analysis of the logical relations between symmetric causality, open time, extrinsic *vs.* intrinsic seriality of time, and anisotropy of time requires us to reject the following statement by Reichenbach:

In the usual discussions of problems of time it has become customary to argue that only irreversible processes supply an asymmetrical relation of causality while reversible processes allegedly lead to a symmetrical causal relation. This conception is incorrect. Irreversible processes alone can define a direction of time; but reversible processes define at least an [serial] order of time, and thereby supply an asymmetrical relation of causality. The reader is referred to the discussion of the relation *to the left of* (. . .). The correct formulation is that only irreversible processes define a *unidirectional* causality.⁷

Reichenbach notes that while the causal processes of classical mechanics and special relativity are reversible, these "reversible" theories nonetheless employ a temporal order which is *serial*. He then infers that (a) even in a reversible world the causal relation must be asymmetric, and (b) in an irreversible world, we require a temporal relation which is not "merely" serial but also "unidirectional," as well as a causal relation which is both asymmetric and unidirectional. But Reichenbach seems to have overlooked completely that if a physical theory affirms the seriality of the time of a completely reversible world, then that seriality is extrinsic and the assignment of the lower of two real numbers as the temporal name to one of two causally connected events therefore does not express any objective asymmetry on the part of the causal relation itself.

⁷ H. Reichenbach: *DT*, *op. cit.*, p. 32.

We are now ready to consider in detail whether thermodynamic processes furnish a physical basis for an anisotropy of time. More specifically, our problem is whether entropy, whose values are given by real numbers, succeeds, unlike the causality of reversible processes, in conferring anisotropy on open time by *intrinsically* defining a serial ordering in the class of states of a closed system. At first, we shall deal with this problem in the light of classical phenomenological thermodynamics. A separate treatment based on the entropy of statistical mechanics will then follow. The second part of this chapter will then consider the existence of *non-entropic* physical bases for the anisotropy of time. It will turn out that both the thermodynamic and the non-thermodynamic species of irreversibility which we shall find to obtain are *de facto* or nomologically-contingent in character.

I. *The Entropy Law of Classical Thermodynamics.*

Suppose that a physical system is created such that one end is hot and the other cold and is then essentially closed off from the rest of the world. In ordinary experience, we do not find that the hot end becomes hotter at the expense of the increased coolness of the cool end. Instead, the system tends towards an equilibrium state of intermediate temperature. And, at least as far as ordinary physical experience is concerned, this entire process of temperature equalization is irreversible. It is possible to characterize this irreversibility more precisely by associating with each momentary state of the closed system a certain quantity, called the "entropy." For the entropy provides the following relative measure of the degree of temperature-equalization attained by the system in the given state: the irreversible temperature-equalization associated with the transition from the initial to the final state corresponds to an increase in the entropy. Accordingly, for a closed system not already in equilibrium, the original *non-statistical* form of the second law of thermodynamics affirms an increase of the entropy with time.⁸

⁸ Although we shall have occasion to see that Clausius's phenomenological principle of entropy increase does indeed require emendation in the light of the discoveries of statistical mechanics, I must dissent from the following judgment by H. Mehlberg ["Physical Laws and Time's Arrow," in H. Feigl and G. Maxwell (eds.) *Current Issues in the Philosophy of Science*, *op. cit.*, p. 115]: "the only rigorous axiomatization of phenomenological thermodynamics (due to Carathéodory) [C. Carathéodory: "Untersuchungen

We shall find the *non*-statistical form of this second law to be untenable in the light of statistical mechanics. It will nonetheless prove useful initially to take Clausius's original form of the law

über die Grundlagen der Thermodynamik," *Mathematische Annalen*, 1909] has also stripped the second phenomenological principle of thermodynamics of its irreversible and anisotropic implications. This important result has been clarified by Professor A. Landé in his illuminating presentation of Carathéodory's contribution [A. Landé: "Axiomatische Begründung der Thermodynamik durch Carathéodory," *Handbuch der Physik*, Vol. IX (1926), pp. 281-300]. But upon turning to the latter article, we find Professor Landé writing as follows: "The theory of quasi-static changes of state (existence of the entropy, etc.) is independent, however, of assertions that might be made about the behavior in the case of non-static processes; thus, for example, all theorems . . . for quasi-static processes would remain valid without change, even if in the case of non-static processes . . . irreversible processes were to run their course in a direction opposite to the actual one." [A. Landé, *ibid.*, p. 299] . . . "Hence our first conclusion is that it is irrelevant for the existence of the entropy whether the quasi-static processes themselves are reversible or not." [A. Landé, *ibid.*, p. 300, quoting from T. Ehrenfest]. "The second conclusion is that the existence of the entropy (or the quasi-static-adiabatic unattainability of neighboring points) is likewise independent of whether the non-static processes are reversible or not. . . . Thus the second law for *quasi-static* changes of state would not even be placed in jeopardy in the event that one were to succeed in making non-static processes time-reversible. To be sure, in that case the principle of Thomson and Clausius for *non-static* processes would become invalid. . . . This case would occur, if one were to include within the scope of one's consideration the temporally inverse processes which, according to the findings of kinetic theory, cannot be excluded." [A. Landé, *ibid.*, p. 300]. It would seem that the Carathéodory-Landé account of *phenomenological* thermodynamics does not justify the claim that Clausius's law for *non-static* processes allows these to be time-reversible. Landé's analysis allows only the following far weaker conclusion, which cannot be adduced as support for Mehlberg's contention: if one does go outside phenomenological thermodynamics and invokes the findings of statistical mechanics (kinetic theory) to affirm the reversibility of *non-static* processes, then one can still uphold the second law of thermodynamics for *quasi-static* processes, although, as Landé points out explicitly, "in that case the principle of . . . Clausius for *non-static* processes would become invalid." Contrary to Mehlberg, it would appear, therefore, that such emendations as must be made in the second phenomenological law of thermodynamics because of its ascription of irreversibility to *non-static* processes do *not* derive at all from the rigor of Carathéodory's axiomatic account of that law within the framework of phenomenological thermodynamics; instead, the required emendations derive wholly from a domain of physical occurrences whose theoretical mastery requires recourse to *statistical* considerations falling outside the purview of phenomenological thermodynamics. Hence, I see no justification for Mehlberg's claim that Carathéodory "has also stripped the second phenomenological principle of thermodynamics of its irreversible and anisotropic implications."

at face value, since A. S. Eddington invoked that version of the law as the basis for attributing anisotropy to time. And although Eddington's account of the anisotropy of time will turn out to lack an adequate physical foundation, a statement of that account and a critique of P. W. Bridgman's misunderstanding of it will be instructive for our subsequent purposes.

The customary statement of the second law of thermodynamics given above has factual content in an obvious sense, *if* the direction of increasing time is defined independently of the entropy-increase. It has been suggested that the independent criterion of time increase be provided either by the continuous matter-energy accretion (as distinct from energy dispersion) postulated by the "new cosmology"⁹ or—in the spatially limited and cosmically brief career of man—by reliance on the subjective sense of temporal order in human consciousness. But I reject both of these proposed criteria. For I am unwilling to base the factual content of so earthy a macroscopic law as that of Clausius on a highly speculative cosmology in which the matter-energy *accretion* manifests itself *macro-empirically* to us merely by the existence of a *steady state*! And we shall see that some important features of man's subjective sense of temporal order can be explained on the basis of the participation of his organism in processes governed by the *entropy statistics* of space ensembles of temporarily closed systems. Would one then have to regard the second law as a tautology, if one were to follow Eddington¹ in resting the anisotropy of physical time on the supposed fact that in one of the two opposite directions of time—which is called the direction of "later"—the entropy of a closed system *increases*, whereas in the opposite direction, the entropy *decreases*? No, this conclusion could not be sustained even if one were to ignore the existence of the viable non-entropic criteria of "later than" to be discussed in Part B of the present chapter. To be sure, if one were to restrict oneself to a *single* closed system and were to say then that of two of its entropy states, the state of greater

⁹ For details on the "new cosmology," see H. Bondi: *Cosmology* (2nd Edition, Cambridge: Cambridge University Press, 1961). A brief digest is given in A. Grünbaum: "Some Highlights of Modern Cosmology and Cosmogony," *The Review of Metaphysics*, Vol. V (1952), pp. 493–98.

¹ Cf. A. S. Eddington: *The Nature of the Physical World*, *op. cit.*, pp. 69ff.

entropy will be called the "later" of the two, then indeed one would have failed to render the factual content of the second, phenomenological law of thermodynamics. But just as in the case of other specifications of empirical indicators — specifications which constitute "definitions" only in the *weak* sense set forth in Chapter One à propos of the "definition" of congruence — Eddington's *non*-statistical entropic "definition" of "later than" was prompted by the presumed *empirical* fact that it does not give rise to contradictions or ambiguities when *different* closed systems are used. For, apart from statistical modifications which we are ignoring for the present, there is concordance in the behavior of *all* closed systems: given any two such systems A and B, neither of which is in thermodynamic equilibrium, if an entropy state S_{A_j} of A is simultaneous with a state S_{B_i} of B, then there is no case of a state S_{A_k} being simultaneous with a state S_{B_l} , such that $S_{A_k} > S_{A_j}$ while $S_{B_l} < S_{B_i}$.² It was this presumed entropic concordance of thermodynamic systems which prompted Eddington's attempt³ to use the second law of thermodynamics in order to account for the anisotropy of time. But by his very unfortunate choice of the misleading name "time's arrow" for the latter feature of physical time, he ironically invited the very misunderstanding which he had been at pains to prevent,⁴ viz., that he was intending to offer a thermodynamic basis for the "unidirectional flow" of psychological time. Eddington maintained that the entropic behavior of closed physical systems distinguishes the two *opposite* directions of time *structurally* in regard to *earlier* and *later* respectively as follows: of two states of the world, the *later* state is the one coinciding with the *higher*

² Examples of other "definitions" which have a corresponding factual foundation in the concordant behavior of different bodies are: first, the "definition" of the metric of time on the basis of the empirical law of inertia, and second, the "definition" of congruence for spatially separated bodies on the basis of the fact that two bodies which are congruent at a given place will be so everywhere, independently of the respective paths along which they are transported individually. Cf. M. Schlick: "Are Natural Laws Conventions," in: H. Feigl and M. Brodbeck (eds.) *Readings in the Philosophy of Science* (New York: Appleton-Century-Crofts, Inc.; 1953), p. 184; H. Reichenbach: "Ziele und Wege der physikalischen Erkenntnis," *Handbuch der Physik*, Vol. IV (1929), pp. 52–53, and *PST*, *op. cit.*, pp. 16–17.

³ A. S. Eddington: *The Nature of the Physical World*, *op. cit.*, pp. 69ff.

⁴ *Ibid.*, pp. 68, 87–110.

entropy of any one closed non-equilibrium system, whereas the *earlier* state corresponds to the state of *lower* entropy. Thus, according to Eddington, the anisotropy of physical time would derive from the supposed fact that in one of the two opposite directions of time which we call the direction of "later"—the entropy of any closed system *increases*, whereas in the opposite direction, the entropy *decreases*.

Since the relation "larger than" for real numbers is serial, Eddington's entropic definition of "later than" intrinsically renders the seriality of time, once its over-all openness is assured by suitable boundary conditions, openness and seriality being attributes concerning which the causal theory of time had to be noncommittal. But, as Eddington seems to have neglected to point out, that theory of time can play an essential role in his entropic "definition" of "later than." For the latter presupposes coordinative "definitions" for the concepts of "temporally between" and "simultaneous" in the very statement of the second law of thermodynamics: this law uses the concept of "closed system," thereby presupposing the concept of "temporal betweenness," as we saw in footnote 9 on page 186 in Chapter Seven, and furthermore the second law requires the concept of simultaneity by making reference to the entropy of an *extended* system *at a certain time*, and implicitly, to the *simultaneous* entropy states of several systems.

P. W. Bridgman has offered a critique of Eddington's entropic "definition" of "later than" which takes as its point of departure the conceptual commitments of the pre-scientific temporal discourse of common sense in which the "Now" of conscious experience is enshrined.⁵ And we shall therefore now find that among the things which render P. W. Bridgman's critique of A. S. Eddington's account of the anisotropy of physical time nugatory, there is Bridgman's erroneous identification of physical time with the "unidirectional flow" of common sense and psychological time.

Bridgman contends that the entropy increase cannot be regarded

⁵ A detailed articulation of the logical commitments of pre-scientific temporal discourse is included in W. Sellars's penetrating "Time and the World Order," *Minnesota Studies in the Philosophy of Science* (Minneapolis: University of Minnesota Press; 1962), Vol. III, pp. 527-616.

with Eddington as a fundamental indicator of the relation "later than." Speaking of the significance which Bridgman believes Eddington to have ascribed to the invariance of the laws of mechanics under time-reversal, Bridgman states his objection to Eddington as follows:

The significance that Eddington ascribes to it is that the equation is unaffected by a reversal of the direction of flow of time, which would mean that the corresponding physical occurrence is the same whether time flows forward or backward, and his thesis is that in general there is nothing in ordinary mechanical occurrences to indicate whether time is flowing forward or backward. In thermodynamic systems on the other hand, in which entropy increases with time, time enters the differential equation as the first derivative, so that the direction of flow of time cannot be changed without changing the equation. This is taken to indicate that in a thermodynamic system time must flow forward, while it might flow backward in a mechanical system.⁶

. . . how would one go to work in any concrete case to decide whether time were flowing forward or backward? If it were found that the entropy of the universe were decreasing, would one say that time was flowing backward, or would one say that it was a law of nature that entropy decreases with time?⁷

We see that Bridgman takes Eddington to have offered an entropic basis of the "forward flow" of psychological time rather than of the anisotropy of physical time, because Bridgman falsely identifies these two different concepts. And we shall now show that his purported *reductio ad absurdum* argument against Eddington's entropic account of the anisotropy of physical time

⁶ P. W. Bridgman: *Reflections of a Physicist* (New York: Philosophical Library; 1950), p. 163.

⁷ *Ibid.*, p. 165. Lest Bridgman's point here be misunderstood, I must point out that in the same essay (pp. 169, 175-77, 181-82) he has explicitly rejected as unfounded the conclusion of statistical mechanics that the entropy of a closed system will *decrease* markedly after being in equilibrium for a very long time. For he rejects as gratuitous the assumption that the micro-constituents of thermodynamic systems can be held to behave reversibly in accord with time-symmetric laws whose observational foundation is only macroscopic. Hence Bridgman's hypothetical finding of an entropy *decrease* is predicated on an actual human observation of an overall entropy decrease with increasing psychological time.

derives its plausibility but also its lack of cogency from the conjunction of precisely this illegitimate identification with a contrary-to-fact assumption. Thus, we ask Bridgman: under what circumstances would it be found that the entropy of the universe "were decreasing"? This situation would arise in the contrary-to-fact eventuality that physical systems or the universe would exhibit *lower* entropy states at times which are *psychologically later*, and *higher* entropy states at times which are *psychologically earlier*, such as in the hypothetical case of finding that lukewarm water separates out into hot and cold portions as time goes on psychologically. To appreciate the import of this contrary-to-fact assumption, we note that an experience *B* is *psychologically later* than an experience *A* under one of the following two conditions: First, the awareness-and-memory-content constituting experience *A* is a *proper part* of the memory-content of experience *B*, or second, experience *B* contains the memory of *the fact* of the occurrence of another experience *A* (e.g., the fact of having dreamt), but the memories ingredient in *B* do not contain the *content* of experience *A* (e.g., the details of the dream having been forgotten).⁸ Thus, psychologically later times are either times at which we do, in fact, have more memories or information than at the correspondingly earlier ones, or they are times at which it would be possible to have a richer store of memories even if the latter did not, in fact, materialize because of partial forgetting. Accordingly, Bridgman's posit of our finding that the entropy "were decreasing" would require the entropy increase among physical systems and the future direction of psychological time to be *temporally counter-directed* as follows: temporally, the direction of increasing entropy among *physical* systems would *not also* be the direction of actual or

⁸ I am indebted to my colleague Professor A. Janis for pointing out to me (by reference to the example of having dreamt) that the first of these conditions is only a sufficient and not also a necessary condition for the obtaining of the relation of being psychologically later. This caution must likewise be applied to the following two assertions by William James, if they are to hold: "our perception of time's flight . . . is due . . . to our memory of a content which it [i.e., time] had a moment previous, and which we feel to agree or disagree with its content now," (W. James: *The Principles of Psychology* [New York: Dover Publications, Inc.; 1950], p. 619), and "what is past, to be known as past, must be known *with* what is present, and *during* the 'present' spot of time." (*Ibid.*, p. 629)

possible *memory* (information) increase among *biological* organisms, since (actually or possibly) "richer" memory states would be coinciding temporally with *lower* entropy states of physical systems.

What is the logical force of Bridgman's contrary-to-fact assumption as a basis for invalidating Eddington's account of the anisotropy of physical time? Bridgman's objection is seen to be devoid of cogency in the light of the following reasons. In the first place, quite apart from the fact that Eddington was not concerned to account for the "forward flow" of psychological time, *in actual fact* the very production of memories in biological organisms depends, as we shall see, on entropy *increases* in certain portions of the external environment. And since Eddington was offering his criterion as an account of what does, in fact, obtain, the adequacy of this account cannot be impugned by the contrary-to-fact logical possibility of counter-directedness envisioned by Bridgman. But even if the situation posited by Bridgman were to materialize in actuality, it would certainly not refute Eddington's claim that (1) the purported entropic behavior of physical systems renders the opposite directions of physical time anisotropic, because the entropy of each of these systems decreases in the one direction and increases in the other, and (2) the direction of increasing entropy can be called the direction of "later than" or time increase. Although Eddington had left himself open to being misunderstood by using the misleading term "time's arrow," he had also sought to spike the misunderstanding that entropically characterized physical time "is flowing forward" in the sense of there being a physical becoming. For he makes a special point of emphasizing⁹ that this becoming, so familiar from *psychological* time, eludes conceptual rendition as an attribute of physical processes, because it involves the concept "now." Contrary to Bridgman, Eddington saw no problem of *physical* time flowing backward rather than forward, since, as Chapter Ten will show in detail, the metaphor "flow" has no relevance to physical time. *Physically*, certain states are later than others by certain amounts of time. But there is no "flow" of *physical* time, because physically there is no egocentric (psychological) transient *now*. Moreover, as applied

⁹ A. S. Eddington: *The Nature of the Physical World*, *op. cit.*, pp. 68-69 and Chapter iv.

to *psychological* time, the locution "flow backward" is self-contradictory, since the assertion that the now shifts forward (in the future direction) is a tautology, as Chapter Ten will demonstrate. A fluid can flow spatially up or down, because the meaning of spatial "flow" is independent of the meaning of "spatially up" or "down." But as applied to psychological time, the meaning of the action-verb metaphor "flow" *involves* the meaning of the metaphor "forward," i.e., of "from earlier to later." For the *flowing* here denotes metaphorically the shifting of the "now" from earlier to later or "forward." Hence if Bridgman's hypothetical situation of counter-directedness could actually materialize, we would say that the entropy is decreasing with increasing psychological time without damage to Eddington's account and not that time is "flowing backward."

Furthermore, if the situation envisioned by Bridgman did arise, we might well not survive long enough to be troubled by it. Poincaré and de Beauregard¹ have explained, in a qualitative way, why prediction and action would probably become impossible under the circumstances posited by Bridgman: two bodies initially at the same temperature would then acquire different temperatures at psychologically later times, while we would be unable to anticipate which of these bodies will become the warmer one, and thus we might be burnt severely if we happen to be in touch with the one that turns out to be the hot one. Or imagine taking a bath in lukewarm water and then not being able to predict which portion of the bathtub will turn out to be the boiling hot end. By the same token, whereas in actuality friction is a *retarding* force, because its dissipation of mechanical energy issues in an increase of the entropy with increasing psychological time, on Bridgman's contrary-to-fact assumption friction would be an *accelerating* force that sets stationary bodies into motion in unpredictable directions. Thus, with increasing psychological time, heat energy would convert itself into mechanical energy of a previously stationary body such as a heavy rock, and prediction of the direction in which the rock would

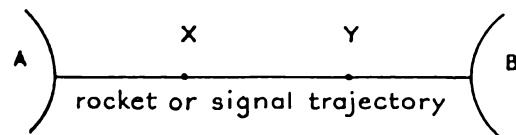
¹ H. Poincaré: *The Foundations of Science*, *op. cit.*, pp. 399–400; O. C. de Beauregard: "L'Irréversibilité Quantique, Phénomène Macroscopique," in A. George (ed.) *Louis de Broglie, Physicien et Penseur* (Paris: Albin Michel; 1953), p. 403, and *Théorie Synthétique de la Relativité Restreinte et des Quanta* (Paris: Gauthier-Villars; 1957), Chapter xiii, esp. pp. 167–71.

start moving would then be well-nigh impossible. And even if we escaped destruction most of the time by not being in the paths of these unpredictable motions, we might well succumb to the anxiety induced by our inability to anticipate and control daily developments in our environment which would constantly threaten our survival.

Finally, suppose that we supplement the hypothetical conditions posited by Bridgman by assuming that in addition to his hypothetical human species *A* whose members are supposed to experience *higher* entropy states of physical systems as psychologically *earlier* than lower ones, there is another human species *B* possessing our actual property of experiencing these same higher entropy states as psychologically later. Then, as Norbert Wiener has noted, a very serious difficulty would arise for communication between our two species *A* and *B* whose psychological time senses are counter-directed. Wiener writes:

It is a very interesting intellectual experiment to make the fantasy of an intelligent being whose time should run the other way to our own. To such a being all communication with us would be impossible. Any signal he might send would reach us with a logical stream of consequents from his point of view, antecedents from ours. These antecedents would already be in our experience, and would have served to us as the natural explanation of his signal, without presupposing an intelligent being to have sent it. . . . Our counterpart would have exactly similar ideas concerning us. *Within any world with which we can communicate, the direction of time is uniform.*²

In amplification of Wiener's statement, consider a situation in which our species *A* and *B* have distinct habitats, which are represented respectively by the regions *A* and *B* of our diagram.



We can then show that any particle or signal which would be regarded as *outgoing* by one of the two species would likewise be held to be *departing* by the other, and any object or message which is *incoming* in the judgment of either species will also be

² N. Wiener: *Cybernetics* (New York: John Wiley & Sons, Inc.; 1948), p. 45. A second edition was published in 1961.

held to be *arriving* by the other. For—to take the case of the *outgoing* influence—suppose that as judged by the members of A, the particle reaches the point Y of its trajectory (see diagram) *later than* the point X and is therefore held to be departing by the men in A. Then the members of B will conclude that the particle is *leaving* them as well, since they will judge that it reaches point X *after* reaching point Y. And thus, if as judged by the A-men, they hurl a rock toward the B region such that the rock comes to rest and remains at rest in B indefinitely, then the B-men, in turn, will judge that a rock, having been at rest in their region all along, suddenly left their habitat and traveled to A, where it was then received by the A-men with ready, open arms. And if the B-men were struck by the discrepancy between the dynamical behavior of that rock and the behavior of other rocks in their habitat—assuming that the *latter* obey the familiar dynamical principles—they might conceivably conclude after a number of such experiences that the dynamically aberrant rocks are linked to the presence elsewhere of temporally counter-directed beings.³ In order to draw this conclusion, however, the

³ Reichenbach (*DT, op. cit.*, pp. 139–40) discusses a situation envisaged by Boltzmann in which there are two entropically counter-directed galactic systems each of which contains intelligent beings whose positive sense of psychological time is geared to the direction of entropy increase in its own galactic environment. Says he (*DT, op. cit.*, p. 139):

Let us assume that among the many galaxies there is one within which time goes in a direction opposite to that of our galaxy. . . . In this situation, some distant part of the universe is on a section of its entropy curve which for us is a downgrade; if, however, there were living beings in that part of the universe, then their environment would for them have all the properties of being on an upgrade of the entropy curve.

And Reichenbach then offers the following quite incomplete hints as to how there might be physical interaction between the two sets of intelligent beings such that either set of beings would be able to secure information indicating the temporal counter-directedness of the other (*DT, op. cit.*, pp. 139–40):

That such a system is developing in the opposite time direction might be discovered by us from some radiation traveling from the system to us and perhaps exhibiting a shift in spectral lines upon arrival . . . the radiation traveling from the system to us would, for the system . . . not leave that system but arrive at it. Perhaps the signal could be interpreted by inhabitants of that system as a message from our system telling them that our system develops in the reverse time direction. We have here a connecting light ray which, for each system, is an arriving light ray annihilated in some absorption process.

B-men would have to assume that the entropy decrease involved in the rock's spontaneous acquisition of kinetic energy from the sand is less probable than the presence elsewhere of temporally counter-directed beings.

Our earlier recognition of the dependence of Eddington's "definition" of "later" on the prior concepts of "temporally between" and "simultaneous" enables us to reply to a further objection raised by Bridgman.⁴

Bridgman claims, on operational grounds, that Eddington's definition is circular and that reliance on the *psychological* sense of what times are "later" is logically indispensable. Says he:

in any operational view of the meaning of natural concepts the notion of time must be used as a primitive concept, which cannot be analyzed, and which can only be accepted, . . . I see no way of formulating the underlying operations without assuming as understood the notion of earlier or later in time.⁵

In an endeavor to show that the specification of the entropy of a closed system at a given instant presupposes the use of the psychological sense of "later," Bridgman says:

[Consider what] is involved in specifying a thermodynamic system. One of the variables is the temperature; it is not sufficient merely to read at a given instant of time an instrument called a thermometer, but there are various precautions to be observed in the use of a thermometer, the most important of

⁴ P. W. Bridgman: *Reflections of a Physicist*, *op. cit.*, pp. 162-67. The rebuttal about to be offered to Bridgman's critique of Eddington's definition also applies to L. Susan Stebbing's arguments against it, as set forth in her *Philosophy and the Physicists* (London: Methuen & Company, Ltd.; 1937), Chapter xi, esp. pp. 262-63.

I should emphasize, however, that my defense of Eddington's "definition" against Bridgman's criticisms must not be construed as agreement with either his general philosophy of science or with his view (cf. *The Nature of the Physical World*, *op. cit.*, pp. 84-85) that the entire universe's supposed past state of *minimum* entropy constitutes a conundrum to which even theological ideas are relevant. For I not only deem theological considerations wholly unilluminating physically in any case (cf. my "Some Highlights of Modern Cosmology and Cosmogony," *The Review of Metaphysics*, *op. cit.*, esp. pp. 497-98, and my "Science and Ideology," *The Scientific Monthly*, Vol. LXXIX [1954], p. 13), but I also maintain that the statistical conception of entropy to be discussed below cuts the ground from under Eddington's assumption that the entire universe must have been *primordially* in a state of *minimum* entropy, an assumption essential to Eddington's puzzlement as to the origin of this presumed state.

⁵ P. W. Bridgman: *Reflections of a Physicist*, *op. cit.*, p. 165.

which is that one must be sure that the thermometer has come to equilibrium with its surroundings and so records the true temperature. *In order to establish this, one has to observe how the readings of the thermometer change as time increases.*⁶

Bridgman claims more than is warranted on precisely the point at issue. For to certify the existence of equilibrium at a certain instant *t*, we must assure ourselves of the absence of a change in the thermometer's reading during a time interval containing the instant *t* other than as an endpoint. But does the procurement of the assurance require a knowledge as to which of the two termini of such an interval is the earlier or later of the two with respect to the usual positive time direction? *Is it not sufficient to ascertain the constancy of the reading between the terminal instants of the time-interval in question?* Indeed, what is presupposed is merely temporal betweenness, which, as we saw, is defined by causal processes, independently of the concept of "later than." But this is hardly damaging to Eddington's "definition" of this concept. And it is irrelevant that, in practice, the experimenter may note also which one of the termini of the time-interval containing the instant *t* is the earlier of the two. For what is at issue is the *semantical* as distinct from the *pragmatic* anchorage of concepts, our inquiry being one in the context of justification and not in the context of discovery.⁷ The complete dispensability of the experimenter's subjective sense of

⁶ *Ibid.*, p. 167, *my italics*.

⁷ For a discussion of Bridgman's unwarranted absorption of semantics within pragmatics, which is a new version of the Sophist doctrine that man is the measure of all things, see A. Grünbaum: "Operationism and Relativity," *The Scientific Monthly*, Vol. LXXIX (1954), pp. 228-31 [reprinted in P. Frank (ed.) *The Validation of Scientific Theories* (Boston: Beacon Press; 1957)]. The same absorption of semantics within pragmatics is found in the following statement by Bridgman ["Reflections on Thermodynamics," *American Scientist*, Vol. XLI (1953), p. 554]: "In general, the meaning of our concepts on the microscopic level is ultimately to be sought in operations on the macroscopic level. The reason is simply that we, for whom the meanings exist, operate on the macroscopic level. The reduction of the meanings of quantum mechanics to the macroscopic level has, I believe, not yet been successfully accomplished and is one of the major tasks ahead of quantum theory."

For a critique of the use of Bridgman's homocentrism in the interpretation of quantum mechanics, see H. Reichenbach: *DT, op. cit.*, p. 224 and A. Grünbaum: "Complementarity in Quantum Physics and Its Philosophical Generalization," *The Journal of Philosophy*, Vol. LIV (1957), p. 719.

earlier and later is further apparent from the fact that the experimenter could certify equilibrium at the instant t , if he were given a film strip showing the constancy of the reading during a time interval between t_1 and t_2 which contains t , *without* being told which end of the film strip corresponds to the earlier moment t_1 . A similar reply can be given to Bridgman's argument⁴ that the *physical* meaning of "velocity" presupposes the psychological sense of earlier and later. For we shall see that purely physical processes in nature define a difference in time-direction quite independently of human consciousness. And thus, for any given choice of the positive space-direction, physical processes themselves define the meanings of both the signs (directions) and magnitudes of velocities independently of man's psychological sense of earlier and later. Here again, Bridgman falsely equates and confuses two *different* meaning components of terms in physics: the physical or semantical with the psychological or pragmatic. The semantical component concerns the properties and relations of purely physical entities which are denoted (named) by terms like "velocity." By contrast, the pragmatic component concerns the activities, both manual and mental, of scientists in *discovering* or coming to *know* the existence of physical entities exhibiting the properties and relations involved in having a certain velocity. That statements about the velocities of masses do not derive their physical meaning from our psychological sense of earlier and later is shown by the fact that cosmogonic hypotheses make reference to the velocities of masses during a stage in the formation of our solar system which preceded the evolution of man and of his psychological time sense. In fact, even in a completely reversible world devoid of beings possessing a time sense, velocity would be a significant attribute of a body despite that hypothetical world's temporal isotropy. But this isotropy would have the consequence that the velocities in such a hypothetical world would not, of course, be anchored in an anisotropic time any more than the positive and negative directions in the presumably isotropic space of our actual world involve the anisotropy of space.

Let us recall Poincaré's and de Beauregard's explanation of why certain sorts of *prediction* and action would become well-

⁴ P. W. Bridgman: *Reflections of a Physicist*, *op. cit.*, p. 167.

nigh impossible under the hypothetical conditions of Bridgman's contrary-to-fact assumption. While still remaining within the framework of the phenomenological second law of thermodynamics, it then becomes apparent that our *actual* world presents us with precisely the inverse temporal asymmetry of inferrability under analogous initial conditions. There are physical conditions in our actual world under which we cannot infer the past but can predict the future. The existence of this particular temporal asymmetry has been obscured by a preoccupation with both reversible processes whose past is as readily determinable as their future, and with conditions in intermittently *open non-equilibrium* systems under which the past can often be inferred from the present, as we shall see presently, whereas the future generally cannot.

Let us clarify the conditions which allow the prediction of the future while precluding the retrodiction of the past by reference to the equation describing a diffusion process, a process in which the entropy increases. This equation is of the form

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = a^2 \frac{\partial \Psi}{\partial t}$$

where a^2 is a real constant. This diffusion equation differs from the wave equation for a reversible process by having a first time derivative instead of a second. In the one-dimensional case of, say, heat-flow, the general solution of the equation governing the temperature Ψ is given by

$$\Psi(x, t) = \sum_{n=1}^{\infty} b_n \sin nx e^{-\frac{n^2}{a^2} t},$$

where the b_n are constants. The behavior of this equation is temporally asymmetric in the following twofold sense: First, if the physical system is in equilibrium at time $t = 0$, then we *cannot* infer what particular sequence of non-equilibrium states issued in the present equilibrium state, since no such sequence is unique,⁹ and second, if the physical system is found in a non-

⁹ This is not to say that there are not other initial conditions under which at least a finite portion of the system's past can be inferred. For a discussion of this case, see J. C. Maxwell: *Theory of Heat* (6th edition; New York: Longmans, Green, and Company; 1880), p. 264, and F. John: "Numerical Solution of the Equation of Heat Conduction for Preceding Times," *Annali di Matematica Pura ed Applicata*, Vol. XL (1955), p. 129.

equilibrium temperature state at $t = 0$, then it could not have been undergoing diffusion for all past values of t , although it can theoretically do so for all future values. Specifically, if external agencies impinge on the system and produce a non-equilibrium state of low entropy at time $t = 0$, then there is no basis for supposing that the system has been undergoing diffusion before $t = 0$, and then the diffusion equation cannot be invoked to infer the "prenatal" past of the system on the basis of its state at $t = 0$, although that equation can be used to predict its future as a closed system undergoing diffusion. Another illustration of the same temporal asymmetry of inferrability in equilibrating processes is given by the case of a ball rolling down the wall on the inside of a round bowl subject to friction: if the ball is found to be at rest at the bottom of the bowl, we *cannot retrodict* its particular motion prior to coming to rest, but if the ball is released at the inside wall near the top, we *can predict* its subsequent coming to rest at the bottom.

This possibility of prophesying the future states of an irreversible process in a closed system under the stipulated conditions in the face of the enigmatic darkness shrouding the non-equilibrium states of the past is so important that E. Hille, following J. Hadamard's analysis of Huyghens's principle in optics, has formulated the fundamental principle of scientific determinism as follows: "From the state of a [closed] physical system at the time t_0 we may deduce its state at a later [but not at an earlier] instant t ."¹

If this be the case, then it is natural to ask why it is that in so many cases involving irreversible processes, we seem to be far more reliably informed concerning the past than concerning the future. This question is raised by Schlick, who points out that human footprints on a beach enable us to infer that a person *was* there in the past but *not* that someone *will* walk there in

¹ E. Hille: *Functional Analysis and Semi-Groups* (New York: American Mathematical Society Publications; 1948) Vol. XXXI, p. 388. Mathematically, the difference between the temporal symmetry of determination in the case of reversible processes and the corresponding asymmetry for irreversible processes expresses itself in the fact that the equations of the former give rise to associated groups of linear transformations while the latter lead to *semi*-groups instead. See also M. S. Watanabe, "Symmetry of Physical Laws. Part III. Prediction and Retrodiction," *Reviews of Modern Physics*, Vol. XXVII (1955), pp. 179-86.

the future. His answer is that "the structure of the past is inferred not from the extent to which energy has been dispersed [i.e., not from the extent of the entropy increase] but from the spatial arrangement of objects."² And he adds that the spatial traces, broadly conceived, are always produced in accord with the entropy principle. Thus, in the case of the beach, the kinetic energy of the person's feet became dispersed in the process of arranging the grains of sand into the form of an imprint, which owes its (relative) persistence in part to the fact that the pedal kinetic energy lost its organization in the course of being imparted to the sand. To be sure, Schlick's claim that the process of leaving a trace occurs in accord with the entropy principle is quite true. But Schlick fails to articulate the logic of the invocation of entropic considerations in the retrodictive inference.

To exhibit the logic of our retrodictive inference that a person did walk on the beach, I briefly anticipate and utilize here results which will emerge from our discussion below of the *statistical* entropy of *temporarily* closed systems. The justification for the inference of the past incursion of the beach by a stroller derives from the following:

For a correction of this overstatement, see Append. §22

(1) most systems which we now encounter in an isolated state of relatively low entropy, behaving as if they might remain isolated, neither were in fact permanently closed in the past nor will remain isolated indefinitely in the future,

(2) in the case of such temporarily isolated or "branch" systems, we can reliably infer a past interaction of the system with an outside agency from a present ordered or low entropy state, an inference which is not feasible, as we shall see in detail, on the basis of the statistical version of the second law of thermodynamics as applied to a single, permanently closed system, and

see Append. §23

(3) the retrodictive inference is based on the assumption that a transition from an earlier high entropy state to a presently given low one is overwhelmingly *improbable* in a system while

² M. Schlick: *Grundzüge der Naturphilosophie* (Vienna: Gerold & Company; 1948), pp. 106–07. J. J. C. Smart ["The Temporal Asymmetry of the World," *Analysis*, Vol. XIV (1954), p. 80] also discusses the significance of traces but reaches the following unwarrantedly agnostic conclusion: "So the asymmetry of the concept of trace has something to do with the idea of formlessness or chaos. But it is not easy to see what." See also Smart's paper in *Australasian Journal of Philosophy*, Vol. XXXIII (1955), p. 124.

it is *isolated*, and this assumption of improbability refers to the frequency of such transitions within a space-ensemble of *branch* systems, each of which is considered at two different times; this improbability does *not* refer to the time-sequence of entropy states of a single, *permanently* closed system.

As applied to the case of the stroller on the beach, these considerations take the following form. We assume that the beach itself was a quasi-closed system not unduly far from equilibrium (smoothness of its sandy surface) for a time in the recent past prior to our encountering it in a foot-shape-bearing state. And we are informed by the discovery of the latter state that the degree of order possessed by the grains of sand is higher and hence the entropy is lower than it should be, in all probability, if the beach had actually remained a quasi-closed system until our present encounter with it. For it is highly improbable that the beach, which is not a permanently closed system, evolved isolatedly from an earlier state of randomness (smoothness) to its present state of greater organization, although the statistical entropy principle for a single permanently closed system does call for precisely such behavior. Hence we conclude that, in all probability, after its initial state of relative smoothness, the beach was an open, interacting system whose increase in order was acquired at the expense of an at least equivalent decrease of organization in the external system with which it interacted (the stroller, who is metabolically depleted). It is clear, therefore, that our retrodictive inference of the stroller's incursion is not made to rest on the premise that in a *permanently* closed system, the entropy never decreases with time, a premise which turns out to be untenable in statistical mechanics, as we shall see presently.

II. *The Statistical Analogue of the Entropy Law.*

Our analysis of entropy so far has been in the macroscopic context of thermodynamics and has taken no adequate account of the important questions which arise concerning the serviceability of the entropy criterion as a basis for the anisotropy of time, when the entropy law is seen in the statistical light of both classical and quantum mechanics. These questions, which we must now face, derive from the attempt to uphold the phenomenological irreversibility of classical thermodynamics in the

face of principles of statistical mechanics asserting that the motions of the microscopic constituents of thermodynamic systems are completely reversible.

If we compare a gas in a highly unequalized state of temperature with a near-equilibrium state from the standpoint of the kinetic theory of gases, we note that the molecular speeds will be much more equalized in the near-equilibrium state of high entropy than in the disequilibrium state of relatively low entropy.

Hence, a *high* entropy corresponds to

- (1) a high degree of molecular equalization
- (2) great homogeneity
- (3) a well-shuffled state
- (4) *low* macro-separation
- (5) low order, where "order" means *not* smoothness or homogeneity, but rather inhomogeneity.

The application of the principles of Newtonian particle mechanics to the constituent molecules of idealized gases takes the following form: each of the n molecules of the gas in the closed system has a position and a velocity, or more accurately, three position coordinates x , y , and z , and three components of velocity. Hence the *micro-state* of the gas can be characterized at any given time by specifying the six position and velocity attributes corresponding to each of the n molecules, each value being given to within a certain small range. The micro-state of the gas at any given time may then be thought of as represented by points in the cells of a six-dimensional position-velocity space or "phase space." And each of the n molecules will then be in some one of the finite number m of cells compatible with the given volume and total energy of the gas.

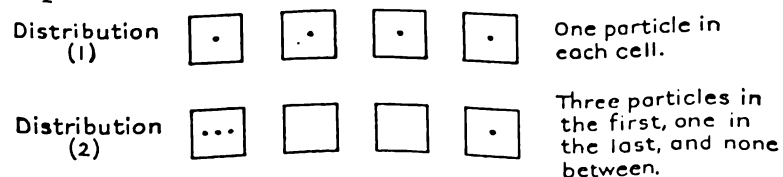
A particular arrangement of the n individual molecules among the m cells constitutes a micro-state of the gas. Thus, if two individual gas molecules A and B were to exchange positions and velocities, a different arrangement would result. However, the *macroscopic* state of the gas, i.e., its being in a state of nearly uniform temperature or very uneven temperature, does not depend on whether it is molecule A or B that occupies a particular point in the container and has a given velocity. What matters *macroscopically* is whether more fast molecules are at one end of the container than at the other end or not, thereby making

the one end hotter, or as hot as the other. In other words, the macro-state depends on *how many* molecules are at certain places in the container, as compared to the number in other places, and also depends on their respective velocities. Thus, the macro-state depends on the numerical spatial and velocity distribution of the molecules, not on the particular *identity* of the molecules having certain positional or velocity attributes. It follows that the *same* macro-state can be constituted by a *number* of *different* micro-states, as in the case of the mere interchange of the microscopic roles played by our two molecules A and B.

It is a basic postulate of statistical mechanics that each one of the m^n possible arrangements or micro-states occur with the same frequency in time or have *the same probability* $1/m^n$. This equi-probability postulate is called the quasi-ergodic hypothesis and gives the so-called probability metric of the Maxwell-Boltzmann statistics, since it asserts what occurrences are equally probable or frequent in time.³

It is of basic importance to see now that the number of micro-states W corresponding to a macro-state of near-equilibrium (uniform temperature) or high entropy is overwhelmingly *greater* than the number corresponding to a disequilibrium state of *non-uniform* temperature or quite low entropy. A drastically oversimplified example will make this fact evident.

Consider a position-velocity or phase-space of only four cells, and let there be just two different distributions (macro-states) of four particles among these cells as follows:



The number of different permutations of four particles in a row is $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$. And thus the number W of different

³ The so-called quasi-ergodic hypothesis is not an assertion based on our lack of knowledge as to the actual relative frequency of the different micro-states: instead it has the logical status of a theoretical claim concerning a presumed fact. What is a matter of our lack of knowledge in this context, however, is which one of the many micro-states that can underlie any given macro-state does, in fact, obtain when the system exhibits the specified macro-state at any one time.

arrangements or micro-states corresponding to the *homogeneous, equalized* macro-state given by distribution (1) is 24. But for the case of the second *inhomogeneous, unequalized* distribution, W is not 24, since the permutations of the three particles *within* the first cell do *not* issue in different arrangements. Hence, for the second case, W has the much smaller numerical value $\frac{4!}{3!} = 4$.⁴

And since the entropy S is given by $S = k \log W$, where k is a constant, the entropy will be lower in the second case than in the first.

It now becomes evident that in the course of time, high entropy states of the gas are enormously more probable or frequent than low ones. For (1) all arrangements are assumed to be equiprobable, i.e., to occur with the same frequency, and (2) many more arrangements correspond to macro-states of high entropy than to states of low entropy. Saying that high entropy states are very probable means that the gas spends the overwhelming portion of its indefinitely long career in the closed system in states of high entropy or equilibrium. This then is the statistical analogue of the law of entropy increase, and it affirms that if a closed system is in a non-equilibrium state of relatively low entropy, then the increase of entropy with time is overwhelmingly probable by virtue of the approach of the particles to their equilibrium distribution. This statistical entropy law is also known as Boltzmann's "H-theorem," the quantity H being related to the entropy S by the relation $S = -kH$, so that an entropy increase corresponds to a decrease in H .

To appreciate the import of this statistical entropy law for the anisotropy of time, we recall that according to Newton's laws, the motions of particles are completely reversible. And all other known laws governing the behavior of the elementary constituents of physical processes likewise affirm the reversibility of

See Append. §25 for additional premisses invoked here

See Append. §25, item (2)

⁴ More generally, Bernoulli's formula for W is $W = \frac{n!}{n_1! n_2! n_3! \dots n_m!}$, where $\sum_1^m n_i = n$. If we wished to *normalize* the thermodynamic probability (which is a large number) so as to be less than 1, then we would have to divide it by the total number of arrangements for all distributions. Hence the (normalized) probability W_p of a *particular* distribution is given by $W_p = \frac{W}{m^n}$.

that behavior. Thus, Maxwell's equations for electromagnetic phenomena, and the fundamental probabilities of state transitions of quantum mechanical systems are time-symmetrical.

Accordingly, we can discuss the case of a gas constituted by Newtonian particles behaving reversibly as the paradigm case for answering the following question: can the statistical form of the entropy law for a permanently closed system form a basis for the anisotropy of time?

Our answer will now turn out to be in the negative. For soon after Boltzmann's enunciation of his theorem, it was felt that there is a logical hiatus in a deduction which derives the overwhelming probability of macroscopic irreversibility from premises attributing complete reversibility to micro-processes. For according to the principle of dynamical reversibility, which is integral to these premises, there is, corresponding to any possible motion of a system, an equally possible reverse motion in which the same values of the coordinates would be reached in the reverse order with reversed values for the velocities.⁵ Thus, since the probability that a molecule has a given velocity is independent of the sign of that velocity, a molecule will have the velocity $+v$ as frequently as the velocity $-v$ in the course of time. And separation processes will occur just as frequently in the course of time as mixing processes, because micro-states issuing in the *unmixing* of hot and cold gases will occur as often as micro-states resulting in their mixing and achieving temperature equalization. J. Loschmidt therefore raised the *reversibility objection* to the effect that for any behavior of a system issuing in an increase of the entropy S with time, it would be equally possible to have an entropy decrease.⁶ Accordingly, the fact that the gas spends most of its career in a state of high entropy does not at all preclude that, in the course of increasing time, the entropy will decrease as often as it increases. A criticism similar to that of Loschmidt was presented in the *periodicity objection*,

⁵ Cf. R. C. Tolman: *The Principles of Statistical Mechanics*, *op. cit.*, pp. 102-04.

⁶ J. Loschmidt: "Über das Wärmegleichgewicht eines Systems von Körpern mit Rücksicht auf die Schwere," *Sitzungsberichte der Akademie der Wissenschaften*, Vienna, Vol. LXXIII (1876), p. 139, and Vol. LXXV (1877), p. 67.

based on a theorem by Poincaré⁷ and formulated by Zermelo.⁸ Poincaré's theorem had led to the conclusion that the long-range behavior of an isolated system consists of a succession of fluctuations in which the value of S will decrease as often as it increases. And Zermelo asked how this result is to be reconciled with Boltzmann's contention that if an isolated system is in a state of low entropy, there is an overwhelming probability that the system is actually in a microscopic state from which changes in the direction of higher values of S will ensue.⁹

These logical difficulties were resolved by the Ehrenfests.¹ They explained that there is no incompatibility between (i) the assertion that *if* the system is in a low entropy state, then, *relative* to that state, it is highly probable that the system will soon be in a higher entropy state, and (ii) the contention that the system plunges down from a state of high entropy to one of lower entropy as frequently as it ascends entropically in the opposite direction, thereby making the *absolute* probability for these two opposite kinds of transition equal. The compatibility of the equality of these two absolute probabilities with a *high* relative probability for a future transition to a higher entropy

⁷ H. Poincaré: "Sur le problème des trois corps et les équations de la dynamique," *Acta Mathematica*, Vol. XIII (1890), p. 67.

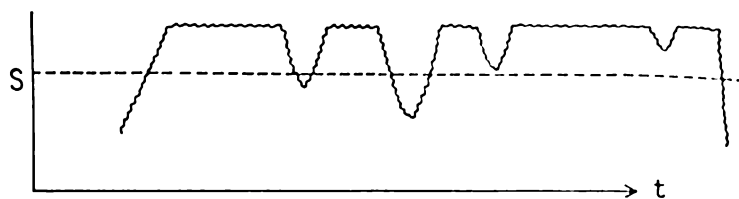
⁸ E. Zermelo: "Über einen Satz der Dynamik und der mechanischen Wärmetheorie," *Wiedemannsche Annalen*, [*Annalen der Physik und Chemie*], Vol. LVII (1896), p. 485.

⁹ For additional details on these objections and references to Boltzmann's replies, see P. Epstein: "Critical Appreciation of Gibbs's Statistical Mechanics," in: A. Haas (ed.) *A Commentary on the Scientific Writings of J. Willard Gibbs* (New Haven: Yale University Press; 1936), Vol. II, pp. 515-19. Cf. also C. Truesdell: "Ergodic Theory in Classical Statistical Mechanics," in: P. Caldirole (ed.) *Ergodic Theories* (New York: Academic Press; 1961), pp. 21-56.

¹ P. and T. Ehrenfest: "Begriffliche Grundlagen der statistischen Auffassung in der Mechanik," *Encyklopädie der mathematischen Wissenschaften*, IV, 2, II, pp. 41-51. See also R. C. Tolman: *The Principles of Statistical Mechanics*, *op. cit.*, pp. 152-58, esp. p. 156; R. Fürth: "Prinzipien der Statistik," in: H. Geiger and K. Scheel (eds.) *Handbuch der Physik* (Berlin: J. Springer; 1929), Vol. IV, pp. 270-72, and H. Reichenbach: "Ziele und Wege der physikalischen Erkenntnis," *op. cit.*, pp. 62-63.

The classical investigations of the Ehrenfests have recently been refined and extended to include quantum theory in D. Ter Haar's important paper "Foundations of Statistical Mechanics," *Reviews of Modern Physics*, Vol. XXVII (1955), pp. 289-338.

becomes quite plausible, when it is remembered that (i) the low entropy states to which the high relative probabilities of subsequent increase are referred are usually at the low point of a trajectory at which changes back to higher values are initiated, and (ii) the Boltzmann H-theorem therefore does not preclude such a system's exhibiting decreases and increases of S with equal frequency. Thus, if we consider a large number of low entropy states of the gas, then we will find that the vast majority of these will soon be *followed* by high entropy states. And in that sense, we can say that it is highly probable that a *low* entropy state will soon be followed by a high one. But it is no less true that a low entropy state was *preceded* by a state of high entropy with equally great probability! The time variation of the entropy, embodying these two claims compatibly, can be visualized as an entropy staircase curve.²



The entropy curve of a permanently closed system.

Boltzmann's H-theorem can thus be upheld in the face of the reversibility and periodicity objections, but only if coupled with a very important proviso: the affirmation of a high probability of a future entropy increase must *not* be construed to assert a *high* probability that low entropy values were *preceded by still lower entropies in the past*. For, as we saw, the relative probability that a *low* entropy state was *preceded* by a state of *higher* entropy is just as great as the relative probability that a low state will be *followed* by a higher state. The fulfillment of the proviso demanded by these results has *two* consequences of fundamental importance whose deduction depends on the statistical entropy law for a *permanently* closed system and which we shall now consider in turn:

1. It destroys the thermodynamic basis for supposing that

² Cf. R. Fürth: "Prinzipien der Statistik," in: H. Geiger and K. Scheel (eds.) *Handbuch der Physik*, *op. cit.*, p. 272.

merely because a present state in a system is one of *low* entropy, this fact is itself sufficient ground for believing that ordered state to be a *veridical trace* of earlier states of still lower entropy which were initiated by a specifiable kind of *interaction* of the system with an outside agency. Von Weizsäcker has rightly suggested³ that, in the absence of other grounds to the contrary, the statistical entropy law itself provides every reason for regarding a present ordered state in a system as a *randomly achieved* low entropy state rather than as a veridical trace of an actual past interaction: it is statistically far more probable in the time sequence of states of a permanently closed system that present low entropy states are mere chance fluctuations rather than the *continuous* successors of actual earlier states of still lower entropy. But present low entropy states in a system cannot serve as articulate documents of the contiguous past of that system, unless we may assume that they directly evolved from specifiable antecedents and hence render veridical testimony of the occurrence of these antecedents in the contiguous past. Hence von Weizsäcker's considerations would suggest that our belief in the *thermodynamic* inferrability of the past is rendered untenable by the verdict of the H-theorem that the entropic behavior of a single, permanently closed system is time-symmetric! Yet it would be a serious error to abandon our ordinary practice of inferring the contiguous past from present low entropy states and to no longer interpret these as traces of interactions with outside agencies. For our supposition that a low entropy state in which we encounter a system is due to the system's past openness or interaction with the outside is based on grounds *other than* the statistically *untenable* assumptions that (i) a system which has been closed for a *very long* time could not now be in a low entropy state or that (ii) the entropy of a permanently closed system never decreases with time. In fact, we shall see that the statistics of the space ensembles of *branch* systems, which we mentioned earlier merely in passing, provide a sound empirical basis for our thermodynamic inferences concerning the past. The existence of this empirical basis, as well

³ C. F. von Weizsäcker: "Der zweite Hauptsatz und der Unterschied von Vergangenheit und Zukunft," *Annalen der Physik*, Vol. XXXVI (1939), p. 281.

as difficulties of its own, undercut the subjectivistic a priori justification of our inferences concerning the past offered by von Weizsäcker on the basis of the alleged transcendental conditions of all possible experience, disclosed in this context by the application of Kant's presuppositional method.

2. Of crucial importance to whether there is an *entropic* basis for the anisotropy of time is the fact that the time-symmetric behavior of the entropy of a permanently closed system would seem to suggest a negative answer to this question. For if we confine ourselves to a *permanently* closed system, statistical mechanics yields the following fundamental results: first, it is impossible to say with Eddington that the *lower* of two given entropy states is the *earlier* of the two, a low entropy state being *preceded* by a high state no less frequently than it is *followed* by a high state,⁴ and second, there is no contradiction between the high relative probabilities of Boltzmann's H-theorem and the equality among the absolute probabilities of the reversibility and periodicity objections, but the aforementioned time-symmetry of the statistical results on which these objections are based shows decisively that *no pervasive anisotropy could possibly be conferred on time by the entropic evolution of a single, permanently closed system*. Hence even if the entire universe qualified as a system whose entropy is defined—which a spatially infinite universe does not, as we shall see later in this chapter—its entropic behavior could not confer any pervasive anisotropy on time. An attempt to find a thermodynamic basis for the anisotropy of time has been made by Max Born by his refusal to affirm the reversibility of elementary processes.⁵ Noting that Boltzmann's averag-

⁴ Cf. H. Reichenbach: *DT*, *op. cit.*, pp. 116–17. It will be recalled that a system can be held to be in the *same macro-state* at different times *t* and *t'* and hence to have the same entropy at *t* and *t'* while the underlying *micro-states* at these times are *different*.

⁵ M. Born: *Natural Philosophy of Cause and Chance* (Oxford: Oxford University Press; 1949), pp. 59, 71–73, and 109–14. In G. Bergmann's review of this work (*Philosophy of Science*, Vol. XVII [1950]), the latter remarks (p. 198) that Born's point can be rendered more clearly by the statement that "in a very relevant sense of the terms statistical mechanics is not mechanics. If, in applying it to, say, the gas, one predicts from a distribution the probability of other distributions, one abandons the idea of orbits and, therefore, deals with 'particles' only in the attenuated sense of using a theory whose fundamental entities have the formal properties of position-momentum coordinates."

In a different vein, L. L. Whyte has since suggested that the reversibility

ing is the expression of our ignorance of the actual microscopic situation, he maintains that the reversibility of mechanics is supplanted by the irreversibility of thermodynamics as a result of "a deliberate renunciation of the demand that in principle the fate of every single particle be determined. You must violate mechanics in order to obtain a result in obvious contradiction to it." He therefore finds that "the statistical foundation of thermodynamics is quite satisfactory even on the basis of classical mechanics."⁶ But it is precisely in the domain of elementary processes that classical mechanics must be superseded by quantum theory. Born therefore attempts to solve the problem by asserting that the new theory "has accepted partial ignorance already on a lower level and need not doctor the final laws" and then offers a derivation of Boltzmann's H-theorem from quantum mechanical principles.⁷

Schrödinger has commented on Born's attempt and has offered an alternative to it. I shall therefore defer considering the status of irreversibility in quantum mechanics until after giving attention to Schrödinger's views.⁸ The assessment of the bearing of quantum mechanics on irreversibility will then be followed by an account of what I consider to be a viable thermodynamic basis for a statistical anisotropy of time.

Referring to Born's account of irreversibility, Schrödinger says:

of elementary processes may have to be abandoned in future physical theory. Says he: "We should give up the long struggle with the question: 'How does irreversibility arise if the basic laws are reversible?' and ask instead: 'If the laws are of a one-way character, under what . . . conditions can reversible expressions provide a useful approximation?'" ("One-way Processes in Physics and Biophysics," *British Journal for the Philosophy of Science*, Vol. VI [1955], p. 110.) The successful implementation of Whyte's proposal would readily provide a previously unsuspected physical basis for the anisotropy of time. But it must not be overlooked that there is such confirmation for fundamental reversibility as the fact that the experimentally substantiated reciprocity law is deducible, as Onsager showed, from the reversibility of elementary collisions. Cf. J. M. Blatt: "Time Reversal," *Scientific American*, August, 1956, pp. 107-14 and R. G. Sachs: "Can the Direction of Flow of Time Be Determined?" *Science*, Vol. CXL (1963), pp. 1284-90.

⁶ M. Born: *Natural Philosophy of Cause and Chance*, *op. cit.*, pp. 72-73.

⁷ *Ibid.*, pp. 110, 113-14.

⁸ E. Schrödinger: "Irreversibility," *Proceedings of the Royal Irish Academy*, Vol. LIII (1950), Sec. A, p. 189, and "The Spirit of Science," in: J. Campbell (ed.) *Spirit and Nature* (New York: Pantheon Books; 1954), pp. 337-41.

"to my mind, in this case, as in a few others, the 'new doctrine' which sprang up in 1925/26 has obscured minds more than it has enlightened them."⁹ His proposal to deal with the issue without a "philosophical loan from quantum mechanics" does not take the form of deriving the increase of entropy with time from some kind of general reversible model. He rejects that approach on the grounds that he is unable to devise a model sufficiently general to cover all physical situations and also suitable for incorporation in all future theories. Neither does he wish to confine himself to a refutation of the arguments directed against Boltzmann's particular reversible model of a gas whose macro-behavior is irreversible.¹ Instead of *deriving* irreversibility, Schrödinger offers to "reformulate the laws of phenomenological irreversibility, thus certain statements of thermodynamics, in such a way, that the logical contradiction *any* derivation of these laws from reversible models seems to involve is removed once and for ever."²

To implement this program, he makes use of the fact that if, during a period of over-all entropy increase or decrease, a system has separated into two subsystems that are isolated from one another, then the respective entropies of the latter will either increase monotonically in both of them (apart from small fluctuations) or will so decrease in both of them. And, instead of considering merely a single isolated system, he envisages at least two systems, called "1" and "2," *temporarily isolated from the remaining universe for a period not greatly exceeding the age of our present galactic system*. Specifically, using a time variable t whose relation to phenomenological time will become clear below, he assumes that systems 1 and 2 are isolated from one another between the moments t_A and t_B , where $t_B > t_A$, but in contact for $t < t_A$ and $t > t_B$. Denoting the entropy of system 1 at time t_A by " S_{1A} " and similarly for the other entropy states, Schrödinger then formulates the entropy law as

$$(S_{1B} - S_{1A})(S_{2B} - S_{2A}) \geq 0.$$

⁹ E. Schrödinger: "Irreversibility," *op. cit.*, p. 189.

¹ Born (*Natural Philosophy of Cause and Chance*, *op. cit.*, p. 59) points out that it was not until recently that the H-theorem was proven for cases other than Boltzmann's model of a gas.

² E. Schrödinger: "Irreversibility," *op. cit.*, p. 191.

Since the law is always applied to those *pairs* of systems having a common branching origin, the product of the entropy differences in it will yield the arithmetical "arrow" of the inequality even in the case of negative entropy differences.

Can Born's quantum mechanical approach or Schrödinger's alternative to it furnish a basis for the anisotropy of time? We saw that Born is guided by the view that since probability enters in quantum mechanics in a fundamental way *ab initio*, the derivation of the probabilistic macroscopic irreversibility affirmed by the H-theorem is feasible in that discipline and not liable to the charge, leveled against Boltzmann's classical derivation, that the deduction depended upon the addition of extraneous probability assumptions to the reversible dynamical equations. But Born's argument is open to important criticisms. To state these, we note first the requirements constituting the quantum mechanical analogue of the classical conditions for the *reversal* of the motion of a closed system: a system N can be said to behave in a manner reverse to that of a system M, if at time *t* it exhibits the same probability for specified values of the coordinates, the same probability for specified values of the momenta taken with reversed sign, and the same expectation value for any function of the coordinates and reversed momenta as would be exhibited by system M at time $-t$. Now, it has been shown by reference to the Schrödinger equation governing the change of isolated (conservative) quantum mechanical systems with time that all three of these conditions are satisfied by such systems.³ The Schrödinger equation for a single free particle relevant here is of the form

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = \frac{4\pi m}{i h} \frac{\partial \Psi}{\partial t}, \text{ where } i \equiv \sqrt{-1}$$

and thus would *seem* to belong formally to the same class as the diffusion equation, which we considered earlier. Due to the presence, however, of an imaginary constant in the Schrödinger equation in place of the real constant in the diffusion equation, it turns out⁴ that the Schrödinger equation describes a *reversible*

³ Cf. R. C. Tolman: *The Principles of Statistical Mechanics*, *op. cit.*, pp. 396-99, and H. Reichenbach: *DT*, *op. cit.*, pp. 207-11.

⁴ See A. Sommerfeld: *Partial Differential Equations in Physics*, trans. E. G. Straus (New York: Academic Press, Inc., 1949), pp. 34-35.

oscillation while the diffusion equation describes an *irreversible* equalization.

What is the physical meaning of the purely formal reversibility of the Schrödinger equation? Instead of being confronted with the classical reversibility of the elementary processes themselves, we now have a *two-wayness* of the transitions between two sets of *probability* distributions of measurable quantities as follows: if nature permits a system which is characterized by the state function Ψ' and the associated set s' of probability distributions at time t_1 to evolve so as to acquire the state function Ψ'' and the associated set s'' of probability distributions at time t_2 , then it also permits the inverse transition from s'' at time t_2 to s' at time t_1 .⁵ S. Watanabe was therefore able to demonstrate that Born's deduction of a monotonic entropy increase with time from the basic principles of quantum mechanics is just as vulnerable to Loschmidt's reversibility objection as the corresponding classical derivation.⁶ And the resulting irrelevance of Born's invocation of the non-deterministic character of the fundamental principles of quantum mechanics is now apparent from the following lucid statement by L. Rosenfeld, who writes:

The introduction of the quantal description of the elementary constituents as a basic assumption instead of the classical picture does not make the least difference to the fundamental structure of statistical thermodynamics; for the quantal laws, just as the classical ones, are reversible with respect to time, and the problem of establishing the macroscopic irreversibility by taking account of the statistical element involved in the concept of macroscopic observation remains unchanged and is again solved by ergodic theorems. The issue has been obscured by the fact that quantum theory itself, in contrast to classical theory, introduces a statistical element at the microscopic level; and it has

⁵ Cf. O. Costa de Beauregard: "Complémentarité et Relativité," *Revue Philosophique*, Vol. CXLV (1955), pp. 397-400.

⁶ S. Watanabe: "Réversibilité contre Irréversibilité en Physique Quantique," *Louis de Broglie, Physicien et Penseur* (Paris: Albin Michel; 1953), p. 393. Cf. also that author's earlier *Le Deuxième Théorème de la Thermodynamique et la Mécanique Ondulatoire* (Paris: Hermann & Cie.; 1935), esp. Chapter iv, Sec. 3, where he shows that, like Newtonian mechanics, quantum mechanics can furnish an irreversible thermodynamics only by adding a distinctly statistical supplementary postulate to its fundamental dynamical principles.

sometimes been confusedly argued that it is this elementary quantal statistics which provides the basis of macroscopic irreversibility. In reality, we have here two completely distinct statistical features, which are not only logically independent of each other, but also without physical influence upon each other. The question whether the elementary law of change is deterministic (as in classical physics) or statistical (as in quantum theory) is entirely irrelevant for the validity of the ergodic theorems.⁷

It will be noted that in articulating the physical meaning of the formal reversibility of the Schrödinger time-equation, we spoke only of two-way transitions *from present to future* states and made no statement concerning inferences from a present state regarding the values we would have obtained in hypothetical *past* measurements, if we had carried them out earlier. There is a very important reason for this deliberate omission, and this reason is the source of a lack of isomorphism between classical reversibility and its quantum mechanical analogue: according to the orthodox version of quantum mechanics, the interaction between the system under observation and the measuring device discontinuously and irreversibly *changes* the Ψ -function characterizing the system before the measurement by imposing a random phase factor on that earlier Ψ -function, and this discontinuous change in Ψ is not governed by Schrödinger's equation. Thus, when the quantum mechanical system is subjected to observation by being coupled indivisibly to a classically describable macroscopic system, a present state function obtained by a measurement of one of the eigenvalues of an observable may be utilized in Schrödinger's equation to determine future but not past values of Ψ . Hence the irreversible alteration of the Ψ -function prevailing before the measurement by the act of measurement, i.e., the *irreversible* changes which take place both in the observed physical system and in the macroscopic measuring apparatus while the latter secures observational information, enter integrally into the (orthodox version of) quantum theory

⁷ L. Rosenfeld: "On the Foundations of Statistical Thermodynamics," *Acta Physica Polonica*, Vol. XIV (1955), p. 9, and "Questions of Irreversibility and Ergodicity," in: P. Caldivola (ed.) *Ergodic Theories* (New York: Academic Press; 1961), pp. 1-20. Cf. also G. Ludwig: "Zum Ergodensatz und zum Begriff der makroskopischen Observablen. I," *Zeitschrift für Physik*, Vol. CL (1958), p. 346.

in marked contrast to classical mechanics and electrodynamics.⁸

We can now understand A. Landé's argument that if we construe reversibility to mean that there are temporal "mirror-images" of physical processes such that the original process and its inverse each comprise an initial state, intermediate states, and a final state, then it is *incorrect* to suppose that the Schrödinger time-equation warrants the ascription of reversibility to elementary quantum mechanical processes for the following reasons: First, actual states are ascertained by particular tests (e.g., states of energy or position), whereas Ψ is not a state but a *statistical link* between two states, and second, the Schrödinger time equation "does not describe processes from an initial to a final state via intermediate [measured] states actually passed through."⁹ But Landé notes elsewhere¹ that the *metrogenic* entropy increase of quantum mechanics is only statistical in the sense that "In reality, the entropy values yielded by successive tests will oscil-

⁸ Details on metrogenic irreversibility in quantum mechanics are given in J. von Neumann: *Mathematische Grundlagen der Quantenmechanik* (Berlin: J. Springer; 1932), pp. 191, 202-12; English trans. R. T. Beyer (Princeton: Princeton University Press; 1955). Cf. also S. Watanabe: "Prediction and Retrodiction," *op. cit.*, p. 179; Watanabe's essay for the *de Broglie Festschrift*: "Réversibilité contre Irréversibilité en Physique Quantique," *op. cit.*, p. 389; D. Bohm: *Quantum Theory* (New York: Prentice-Hall; 1951), Chapter xxii, and S. Watanabe: "Le Concept de Temps en Physique Moderne et la Durée Pure de Bergson," *Revue de Métaphysique et de Morale*, Vol. LVI (1951), pp. 134-35. Reichenbach does not take cognizance of quantum mechanical metrogenic irreversibility in his theory of the direction of time (cf. *DT*, *op. cit.*, Chapter xxiv, and "Les Fondements Logiques de la Mécanique des Quanta," *Annales de l'Institut Henri Poincaré*, Vol. XIII [1953], pp. 148-54).

In an article "Philosophical Problems Concerning the Meaning of Measurement in Physics" (*Philosophy of Science*, Vol. XXV [1958]), H. Margenau has contested the "orthodox" conception of the process of measurement as the reduction of a wave packet. And thus he rejects the necessity of associating with the process of measurement a discontinuous change in the ψ -function that is not governed by the Schrödinger equation. On this "unorthodox" view, the claims made above on the basis of the "orthodox" version would have to be revised accordingly. Cf. also H. Margenau: "Measurements and Quantum States," *Philosophy of Science*, Vol. XXX (1963), pp. 1-16.

⁹ A. Landé: "The Logic of Quanta," *British Journal for the Philosophy of Science*, Vol. VI (1956), p. 300, esp. pp. 305-07 and 311.

¹ A. Landé: "Wellenmechanik und Irreversibilität," *Physikalische Blätter*, Vol. XIII (1957), pp. 312-14. For a general discussion of these issues, see M. M. Yanase: "Reversibilität und Irreversibilität in der Physik," *Annals of the Japan Association for Philosophy of Science*, Vol. I (1957), pp. 131-49.

late up and down just as the classical entropy values of the Ehrenfest curve."

A brief comment concerning the epistemological status of metrogenic irreversibility in quantum mechanics must precede our assessment of its capabilities to account for the anisotropy of our macro-time.

Guided by the precepts of philosophical idealism, Watanabe erroneously equates the observer qua recorder of physically registered observational data with the observer qua *conscious* organism. He then infers that the metrogenic irreversibility of quantum mechanics shows "decisively" that "there is no privileged direction in the time of physics, and that, if one finds a unique direction in the evolution of physical phenomena, this is merely the projection of the flow of our psychic time . . . the increase in entropy is not a property of the external world left to itself, but is the result of the union of the subject and the object."² Treating the seriality inherent in psychological time as autonomous and *sui generis*, he nevertheless admits that the uniformity of psychological time-directions as between different living organisms requires explanation, being too remarkable to be contingent. But he seeks the explanation along the lines of Bergson's very questionable conception that living processes obey autonomous principles.³ In his mentalistic interpretation of metrogenic irreversibility in quantum mechanics, Watanabe requires, as a crucial tacit premise in his argument, the traditional idealist characterization of the status of such material common sense objects as the classically describable pieces of apparatus used, in one way or another, in all quantum mechanical measure-

² S. Watanabe: "Le Concept de Temps en Physique Moderne et la Durée Pure de Bergson," *op. cit.*, pp. 134-36. Cf. also that author's contribution to the de Broglie *Festschrift*: "Réversibilité contre Irréversibilité en Physique Quantique," *op. cit.*, pp. 385, 392, 394.

³ For a detailed discussion of the role of *physical* irreversibility in biological processes, see H. F. Blum: *Time's Arrow and Evolution* (2nd edition; Princeton: Princeton University Press; 1955); E. Schrödinger: *What is Life?* (New York: The Macmillan Company; 1945), Chapter vi; and R. O. Davies: "Irreversible Changes: New Thermodynamics from Old," *Science News* (May, 1953), No. 28. Attempts to prove the autonomy of living processes can draw no support from instances of entropy decrease in the human body. For, being an open system, that body's entropy can decrease or increase in complete conformity with even the *non*-statistical second law of thermodynamics.

ments. But this idealist premise is altogether unconvincing, and without it, there is every reason to regard the interaction between physical systems and the observational devices used in quantum mechanics as an entirely physical matter devoid of psychological ingredients of any kind. For, as has been explained by von Neumann⁴ and more recently, by Ludwig,⁵ the demand that cognizance be taken of the disturbances produced by measurements and observation can be adequately met in quantum mechanics without including the human observer's retina or body in the analysis, let alone his stream of consciousness. In regard to the macroscopic system which undergoes irreversible changes in the course of registering the results of microphysical measurements, Ludwig points out that, in principle, the perception of its readings by a conscious subject is irrelevant. Says he: "in principle, it is not necessary that it was a physicist [i.e., human observer] who built the apparatus for the purpose of measurement. It can also be a system on which the microscopic object impinges, entirely in the natural course of events."⁶ Thus, as far as the role of the human observer *qua* conscious organism is concerned, there is no epistemological difference between quantum mechanics and classical physics.

Although quantum irreversibility is an entirely physical matter and a quantum world precludes our ascribing certain kinds of physical properties to a physical system apart from the interaction of that system with specifiable measuring devices, the irreversibility of our ordinary environment cannot be held to be attributable to metrogenic quantum irreversibility alone. Bohr's principle of complementarity must be taken in conjunction with his own emphasis in the correspondence principle that the meas-

⁴ J. von Neumann: *Mathematische Grundlagen der Quantenmechanik*, *op. cit.*, pp. 187, 223-37, esp. pp. 223-24. Cf. also D. Bohm: *Quantum Theory*, *op. cit.*, pp. 584-85, 587-90, 600-09.

⁵ G. Ludwig: "Der Messprozess," *Zeitschrift für Physik*, Vol. CXXXV (1953), p. 483, esp. p. 486; see also his *Die Grundlagen der Quantenmechanik* (Berlin: J. Springer; 1954), pp. 142-59, 178-82, and his "Die Stellung des Subjekts in der Quantentheorie," in *Veritas, Justitia, Libertas*, Festschrift zur 200-Jahr Feier der Columbia University (Berlin: Colloquium Verlag; 1954), pp. 261-71. See also H. Reichenbach: *DT*, *op. cit.*, pp. 223-24, and *Philosophic Foundations of Quantum Mechanics* (Berkeley: University of California Press; 1948), pp. 15ff.

⁶ G. Ludwig, "Der Messprozess," *Zeitschrift für Physik*, *op. cit.*, p. 486.

uring devices which constitute the epistemological basis of quantum mechanics are themselves describable by the principles of classical physics. The actual irreversibility of our macro-environment is set in a context in which Planck's constant h may be considered negligibly small and in which the classical view that the physical system can be said to have definite physical properties independently of any coupling to a measuring instrument is legitimately applicable.⁷

We can therefore endorse Schrödinger's *rejection* of the use of quantum mechanical metrogenic irreversibility as a basis for explaining the phenomenological (macro-) irreversibility of our environment. Says he: "Surely the system continues to exist and to behave, to undergo irreversible changes and to increase its entropy in the interval between two observations. The observations we might have made in between cannot be essential in determining its course."⁸

Granted then that quantum mechanics does not furnish the required account of the temporal anisotropy of our macroscopic environment in its "current" non-equilibrium state, does Schrödinger's own non-quantal account succeed in doing so? He avowedly made no attempt to deduce irreversibility. But he does explain that if at least one of the entropy differences in his formulation of Clausius's principle is positive, then it is the parametric time t which corresponds to phenomenological time and that, alternatively, if at least one such difference is negative, it is $-t$ that corresponds to phenomenological time. Schrödinger's perceptive guiding idea that the attempt to characterize the phenomenological anisotropy of time entropically without run-

⁷ For an interesting discussion of the conditions governing such applicability, see L. Brillouin: *Science and Information Theory* (New York: Academic Press; 1956), pp. 229-32.

⁸ E. Schrödinger: "Irreversibility," *op. cit.*, p. 190. Even less helpful than quantum irreversibility as a basis for the anisotropy of time is the Wheeler-Feynman-Stückelberg analysis of pair production in quantum electrodynamics (cf. H. Margenau: "Can Time Flow Backwards?" *Philosophy of Science*, Vol. XXI [1954], p. 79), since it involves indeterminacies even in regard to those order properties of time which are defined by reversible macro-processes (cf. H. Reichenbach: *DT*, *op. cit.*, pp. 262-69, and "Les Fondements Logiques de la Mécanique des Quanta," *op. cit.*, pp. 150-53). See also C. W. Berenda: "Determination of Past by Future Events," *Philosophy of Science*, Vol. XIV (1947), p. 13.

ning afoul of the reversibility and periodicity objections can succeed only if we regard the entropy law as an assertion about *at least two temporarily* closed systems was developed independently by Reichenbach. And the valid core—but only the valid core—of Reichenbach's version of this idea seems to me to provide a foundation for an entropic basis of a statistical anisotropy of physical time. Believing that Reichenbach's account requires significant modification in order to be satisfactory, I shall now set forth what I consider to be a corrected elaboration of his principal conception.

Actual physical experience presents us with entropy increases in quasi-isolated systems overwhelmingly more frequently than with corresponding decreases: if 10,000 people sat down together to dinner and each poured some cream into a cup of black coffee, it is an incontestably safe bet that the cream will mix with the coffee in all cases and that no one will report a subsequent *unmixing* of them *for ordinary intervals of time*, i.e., prior to the consumption of the creamed coffee. This kind of phenomenon of temporally asymmetric entropy increase is, of course, not incompatible with the statistical form of the entropy law for a permanently closed system, since we restricted ourselves to ordinary intervals of time. Let us inquire, therefore, whether this kind of phenomenon of entropy increase confers anisotropy at least on the time of our galactic system during the current epoch. We shall now see in detail that it is indeed correct to conclude that such phenomena do furnish a viable physical basis for a statistical anisotropy of time. To do so, we must first describe certain features of the physical world *having the character of initial or boundary conditions* within the framework of the theory of statistical mechanics. The sought-after entropic basis of a statistical anisotropy of time will then emerge from principles of statistical mechanics relevant to these *de facto* conditions.

The universe around us exhibits striking disequilibria of temperature and other inhomogeneities. In fact, we live in virtue of the nuclear conversion of the sun's reserves of hydrogen into helium, which issues in our reception of solar radiation. As the sun dissipates its reserves of hydrogen via the emission of solar radiation, it may heat a terrestrial rock embedded in snow during the daytime. At night, the rock is no longer exposed to the sun but is left with a considerably higher temperature than

the snow surrounding it. Hence, at night, the warm rock and the cold snow form a quasi-isolated subsystem of either our galactic or solar system. And the relatively low entropy of that subsystem was purchased at the expense of the dissipation of the sun's reserves of hydrogen. Hence, *if* there is some quasi-closed system comprising the sun and the earth, the branching off of our subsystem from this wider system in a state of low entropy at sunset involved an entropy increase in the wider system. During the night, the heat of the rock melts the snow, and thus the entropy of the rock-snow system increases. The next morning at sunrise, the rock-snow subsystem merges again with the wider solar system. Thus, there are subsystems which branch off from the wider solar or galactic system in a state of relatively low entropy, remain quasi-closed for a *limited* period of time, and then merge again with the wider system from which they had been separated. Following Reichenbach,⁹ we have been using the term "branch system" to designate this kind of subsystem.

See Append. § 23

Branch systems are formed not only in the natural course of things, but also through human intervention: when an ice cube is placed in a glass of warm gingerale by a waiter and then covered for hygienic purposes, a subsystem has been formed. The prior freezing of the ice cube had involved an entropy increase through the dissipation of electrical energy in some large quasi-closed system of which the electrically run refrigerator is a part. While the ice cube melts in the covered glass subsystem, that quasi-closed system increases its entropy. But it merges again with another system when the then chilled gingerale is consumed by a person. Similarly for a closed room that is closed off and then heated by burning logs.

Thus, our environment abounds in branch systems whose initial relatively low entropies are the products of their earlier coupling or interaction with outside agencies of one kind or another. This rather constant and ubiquitous formation of a branch system in a relatively low entropy state resulting from interaction often proceeds at the expense of an entropy increase in some wider quasi-closed system from which it originated. And the *de facto*, nomologically contingent occurrence of these branch systems has the following *fundamental consequence*, at

⁹ Cf. H. Reichenbach: *DT*, *op. cit.*, pp. 118-43.

least for our region of the universe and during the current epoch: among the quasi-closed systems whose entropy is relatively low and which behave as if they might remain isolated, the vast majority have not been and will not remain permanently closed systems, being branch systems instead.

Hence, upon encountering a quasi-closed system in a state of fairly low entropy, we know the following to be overwhelmingly probable: the system has not been isolated for millions and millions of years and does not just *happen* to be in one of the infrequent but ever-recurring low entropy states exhibited by a permanently isolated system. Instead, our system was formed not too long ago by branching off after an interaction with an outside agency. For example, suppose that an American geologist is wandering in an isolated portion of the Sahara desert in search of an oasis and encounters a portion of the sand in the shape of "Coca Cola." He would then infer that, with overwhelming probability, a kindred person had interacted with the sand in the recent past by tracing "Coca Cola" in it. The geologist would not suppose that he was in the presence of one of those relatively low entropy configurations which are assumed by the sand particles spontaneously but very rarely, if beaten about by winds for millions upon millions of years in a state of effective isolation from the remainder of the world.

There is a further *de facto* property of branch systems that concerns us. For it will turn out to enter into the temporally asymmetrical statistical regularities which we shall find to be exhibited in the entropic behavior of these systems. This property consists in the following *randomness* obtaining *as a matter of nomologically contingent fact* in the distribution of the W_1 micro-states belonging to the initial macro-states of a space-ensemble of branch systems each of which has the same initial entropy $S_1 = k \log W_1$: For each class of *like* branch systems having the *same* initial entropy value S_1 , the micro-states constituting the identical initial macro-states of entropy S_1 are *random samples* of the set of all W_1 micro-states yielding a macro-state of entropy S_1 .¹ This attribute of randomness of micro-states on the part of the initial states of the members of the *space-*

See correction in
Append. §26

¹ Cf. R. C. Tolman: *The Principles of Statistical Mechanics*, *op. cit.*, p. 149.

ensemble will be recognized as the counterpart of the following attribute of the micro-states of one single, permanently closed system: there is equi-probability of occurrence among the W_i micro-states belonging to the *time*-ensemble of states of equal entropy $S_i = k \log W_i$ exhibited by one single, permanently closed system.

We can now state the statistical regularities which obtain as a consequence of the de facto properties of branch systems just set forth, when coupled with the principles of statistical mechanics. These regularities, which will be seen to yield a *temporally asymmetric* behavior of the entropy of *branch* systems, fall into two main groups as follows:²

Group 1. In most space-ensembles of quasi-closed branch systems each of which is initially in a state of non-equilibrium or relatively *low* entropy, the majority of branch systems in the ensemble will have *higher* entropies *after* a given time t . But these branch systems simply did not exist as quasi-closed, distinct systems at a time t *prior to* the occurrence of their initial, branching off states. Hence, not existing then as such, the branch systems did in fact *not* also exhibit the same higher entropy states at the *earlier* times t , which they would indeed have done then had they existed as closed systems all along. In this way, the space-ensembles of branch systems do *not* reproduce the entropic time-symmetry of the single, permanently closed system. And whatever the behavior of the components of the branch systems prior to the latter's "birth," that behavior is irrelevant to the entropic properties of branch systems as such.

The increase after a time t in the entropy of the overwhelming majority of branch systems of initially low entropy—as confirmed abundantly by observation—can be made fully intelligible. To do so, we note the following property of the *time*-ensemble of entropy values belonging to a single, permanently closed system and then affirm that property of the space-ensembles of branch systems: since *large* entropic downgrades or decreases are *far less* probable (frequent) than moderate ones, the *vast majority* of *non-equilibrium* entropy states of a permanently closed system

² Cf. R. Fürth: "Prinzipien der Statistik," *op. cit.*, pp. 270 and 192–93. The next-to-the-last sentence on p. 270 is to be discounted, however, since it is self-contradictory as it stands and incompatible with the remainder of the page.

are located either at or in the immediate temporal vicinity of the *bottom* of a *dip* of the one-system entropy curve. In short, the vast majority of the *sub-maximum* entropy states are on or temporally very near the *upgrades* of the one-system curve. The application of this result to the space-ensemble of branch systems whose initial states exhibit the aforementioned de facto property of *randomness* then yields the following: among the initial low entropy states of these systems, the vast majority lie at or in the immediate temporal vicinity of the bottoms of the one-system entropy curve at which an upgrade begins.

Group 2. A decisive *temporal asymmetry* in the statistics of the temporal evolution of branch systems arises from the further result that in most space ensembles of branch systems each of whose members is initially in a state of *equilibrium* or very *high* entropy, the vast majority of these systems in the ensemble will *not* have *lower* entropies *after* a finite time *t*, but will still be in equilibrium. For the aforementioned randomness property assures that the vast majority of those branch systems whose initial states are equilibrium states have maximum entropy values lying somewhere *well within* the plateau of the one-system entropy curve, rather than at the extremity of the plateau at which an entropy *decrease* is initiated.

Although the decisive asymmetry just noted was admitted by H. Mehlberg,³ he dismisses it as expressing "merely the factual difference between the two relevant values of probability." But an asymmetry is no less an asymmetry for depending on de facto, nomologically contingent boundary conditions rather than being assured by a *law* alone. Since our verification of laws generally has the same partial and indirect character as that of our confirmation of the existence of certain complicated de facto boundary conditions, the assertion of an asymmetry depending on de facto conditions is generally no less reliable than one wholly grounded on a law. Hence when Mehlberg⁴ urges against Schrödinger's claim of asymmetry that for every pair of branch systems which change their entropy in one direction, "there is nothing to prevent" another pair of closed subsystems from changing their entropy in the opposite direction, the reply is:

³ H. Mehlberg: "Physical Laws and Time's Arrow," *op. cit.*, p. 129.

⁴ *Ibid.*, p. 117, n. 30.

Mehlberg's criticism can be upheld only by gratuitously neglecting the statistical asymmetry admitted but then dismissed by him as "merely" factual. For it is the existence of the specified boundary conditions which statistically *prevents* the existence of entropic time-symmetry in this context.

We see therefore that in the vast majority of branch systems, either one end of their finite entropy curves is a point of low entropy and the other a point of high entropy, or they are in equilibrium states at both ends as well as during the intervening interval. And it is likewise apparent that the statistical distribution of these entropy values on the time axis is such that the vast majority of branch systems have the *same direction of entropy increase* and hence also the same opposite direction of entropy decrease. Thus, the statistics of entropy increase among branch systems assure that *in most space ensembles the vast majority* of branch systems will increase their entropy in *one* of the two opposite time directions and decrease it in the other: in contradistinction to the entropic time-symmetry of a single, permanently closed system, the probability within the space-ensemble that a low entropy state *s* at some given instant be followed by a higher entropy state *S* at some given later instant is much *greater* than the probability that *s* be preceded by *S*. In this way *the entropic behavior of branch systems confers the same statistical anisotropy on the vast majority of all those cosmic epochs of time during which the universe exhibits the requisite disequilibrium and contains branch systems satisfying initial conditions of "randomness."*

See
Append. §§ 21 and 51

Let us now call the direction of entropy increase of a *typical representative* of the aforementioned kind of cosmic epoch the direction of "later," as indeed we have done from the outset by the mere assignment of higher time numbers in that direction but *without prejudice* to our findings concerning the issue of the anisotropy of time. Then our results pertaining to the entropic behavior of branch systems show that the directions of "earlier than" and "later than" are not merely opposite directions bearing decreasing and increasing time coordinates respectively but are statistically *anisotropic* in an objective physical sense. For we saw earlier in this chapter that increasing real numbers can be assigned as time coordinates in a physically meaningful way

without any commitment to the existence of (de facto or nomologically) irreversible kinds of processes. In fact, the use of the real number continuum as a basis for coordinatizing time no more entails the anisotropy of time than the corresponding coordinatization of one of the three dimensions of space commits us to the anisotropy of that spatial dimension.

It should be noted that I have characterized the positive direction of time as the direction of entropy increase in branch systems for a typical representative of all those epochs of time during which the universe exhibits the requisite disequilibrium and contains branch systems satisfying initial conditions of "randomness." Accordingly, it is entirely possible to base the customary temporal description of fluctuation phenomena, i.e., the assertion that the entropy *decreases* with *positive time* in some systems or other, on entropic counter-directedness of the latter systems with respect to the *majority* of branch systems, a description which is not liable to the *reductio ad absurdum* offered by Bridgman in his unsuccessful attempt to show that there can be no entropic basis for the anisotropy of time. The very statistics of branch systems which invalidate Bridgman's attempted *reductio* likewise serve to discredit the following denial by K. R. Popper of the relevance of entropy statistics to the anisotropy of time, despite the soundness of Popper's objection to Boltzmann's original formulation:

The suggestion has been made (first by Boltzmann himself) that the arrow of time is, either by its very nature, or by definition, connected with the increase in entropy; so that entropy cannot decrease in time because a decrease would mean a reversal of its arrow, and therefore an increase relative to the reversed arrow. Much as I admire the boldness of this idea, I think that it is absurd, especially in view of the undeniable fact that thermodynamic fluctuations do exist. One would have to assert that, within the spatial region of the fluctuation, all clocks run backwards if seen from outside that region. But this assertion would destroy that very system of dynamics on which the statistical theory is founded. (Moreover, most clocks are non-entropic systems, in the sense that their heat production, far from being essential to their function, is inimical to it.)

I do not believe that Boltzmann would have made his suggestion after 1905, when fluctuations, previously considered no

more than mathematically calculable near-impossibilities, suddenly became the strongest evidence in favour of the physical reality of molecules. (I am alluding to Einstein's theory of Brownian motion.) As it is, a statistical theory of the arrow of time seems to me unacceptable.⁵

I have contended against Mehlberg and Popper that the entropic behavior of branch systems confers the *same* statistical anisotropy on the vast majority of all those cosmic epochs of time during which the universe exhibits the requisite disequilibrium and contains branch systems satisfying the specified initial conditions of "randomness." My claim of statistical anisotropy departs significantly from Reichenbach's "hypothesis of the branch structure"⁶ in the following ways: (1) I do not assume that the entropy is defined for the entire universe such that the universe as a whole can be presumed to exhibit the entropic evolution of the statistical entropy curve for a permanently closed, *finite* system, an assumption which leads Reichenbach to affirm the parallelism of the direction of entropy increase of the universe and of the branch systems at any time, and therefore, (2) I do not conclude, as Reichenbach does, that cosmically the statistical anisotropy of time is only local by "fluctuating" in the following sense: the supposed alternations of epochs of entropy increase and decrease of the universe go hand-in-hand with the alternations of the direction of entropy increase in the ensembles of branch systems associated with these respective epochs, successive disequilibrium epochs allegedly being entropically *counter-directed* with respect to each other.

In view of the reservations which Reichenbach himself expressed⁷ concerning the reliability of assumptions regarding the universe as a whole in the present state of cosmology, one wonders why he invoked the entropy of the universe at all instead of confining himself, as I have done, to the much weaker assumption of the existence of states of disequilibrium in the universe. More fundamentally, it is unclear how Reichenbach thought he could reconcile the assumption that the branch systems satisfy initial conditions of randomness during whatever cosmic epoch

⁵ K. R. Popper: *Nature*, Vol. CLXXXI (1958), p. 402.

⁶ H. Reichenbach: *DT*, *op. cit.*, p. 136.

⁷ *Ibid.*, pp. 132-33.

they may form—an assumption which, as we saw, makes for the *same* statistical anisotropy on the part of *most* disequilibrium epochs of the universe—with the following claim of alternation made by him: “When we come to the downgrade [of the entropy curve of the entire universe], always proceeding in the same direction [along the time-axis], the branches begin at states of high entropy . . . and they end at points of low entropy.”⁸ Contrary to Reichenbach, we saw in our statement of the consequences of the postulate of randomness under *Group 2* above that in the vast majority of cases, branch systems beginning in a state of equilibrium (high entropy) will *remain* in equilibrium for the duration of their finite careers instead of decreasing their entropies!

An inherent limitation on the applicability of the Maxwell-Boltzmann entropy concept to the entire universe lies in the fact that it has no applicability at all to a *spatially infinite* universe for the following reasons. If the infinite universe contains a denumerable *infinity* of atoms, molecules or stars, the number of complexions *W* becomes infinite, so that the entropy is not defined, and a fortiori no increase or decrease thereof.⁹ And if the number of particles in the infinite universe is only finite, then (a) the equilibrium state of maximum entropy cannot be realized by a *finite* number of particles in a phase-space of *infinitely* many cells, since these particles would have to be *uniformly* distributed among these cells, and (b) the quasi-ergodic hypothesis, which provides the essential basis for the probability metric ingredient in the Maxwell-Boltzmann entropy concept, is presumably false for an infinite phase space.¹ If the universe *were finite and* such that an entropy is defined for it as

⁸ *Ibid.*, p. 126.

⁹ Cf. K. P. Stanyukovic: “On the Increase of Entropy in an Infinite Universe,” *Doklady, Akademiia Nauk SSSR*, N.S., Vol. LXIX (1949), p. 793, in Russian, as summarized by L. Tisza in *Mathematical Reviews*, Vol. XII (1951), p. 787.

¹ For additional doubts concerning the cosmological relevance of the entropy concept, cf. E. A. Milne: *Sir James Jeans* (Cambridge: Cambridge University Press; 1952), pp. 164–65, and *Modern Cosmology and the Christian Idea of God* (Oxford: Clarendon Press; 1952), pp. 146–50; also L. Landau and E. Lifschitz: *Statistical Physics* (2nd ed.; Reading: Addison-Wesley; 1958), pp. 22–27.

a whole which conforms to the one-system entropy curve of statistical mechanics, then my contention of a *cosmically pervasive statistical anisotropy* of time could no longer be upheld. For I am assuming that the vast majority of branch systems in most epochs increase their entropy in the same direction and that space ensembles of branch systems do form during most periods of disequilibrium. And if one may further assume that the entropy of a finite, spatially-closed universe depends *additively* on the entropies of its component subsystems, then the assumed temporal asymmetry of the entropy behavior of the branch systems would appear to contradict the complete *time-symmetry* of the one-system entropy behavior of the finite universe. This conclusion, if correct, therefore poses the question—which I merely wish to ask here—whether in a closed universe, the postulate of the randomness of the initial conditions would not hold. For in that case, the cosmically *pervasive* statistical anisotropy of time which is assured by the randomness postulate would not need to obtain; instead, one could then assume initial conditions in branch systems that issue in Reichenbach's cosmically local kind of anisotropy of time, successive overall disequilibrium epochs having *opposite* directions of entropy increase both in the universe and in the branch systems associated with these epochs.

We shall show in Chapter Nine that our account of the entropic basis for the anisotropy of time has the following further important ramifications: First, it provides an empirical justification for interpreting present ordered states as veridical traces of actual past interaction events, a justification which the entropic behavior of a single, permanently closed system was incompetent to furnish, as we saw, and, as will likewise be shown in Chapter Nine, second, it explains why the subjective (psychological) and objective (physical) directions of positive time are parallel to one another by noting that man's own body participates in the entropic lawfulness of space ensembles of physical branch systems in the following sense: man's *memory*, just as much as all purely physical recording devices, accumulates "traces," records or information in a direction dictated by the statistics of physical branch systems. Contrary to Watanabe's

conception of man's psychological time sense as *sui generis*, it will turn out in Chapter Nine that the future direction of psychological time is parallel to that of the accumulation of traces (increasing information) in interacting systems, and hence parallel to the direction defined by the positive entropy increase in the branch systems. Thus the examination of the anisotropy of psychological time will also have shown that Spinoza was in error when he wrote Oldenburg that "*tempus non est affectio rerum sed merus modus cogitandi*."

We have now completed our discussion of the anisotropy of time insofar as it depends on systems for which an entropy is defined and whose entropy changes in a temporally asymmetric way. It remains to consider whether there are not also *non-entropic* kinds of physical processes which contribute to the anisotropy of time. We shall find that the answer is indeed affirmative. Moreover, it will turn out that just as the *entropically* grounded statistical anisotropy of time was *not* assured by the laws alone but also depended on the role played by specified kinds of boundary conditions, so also the *non-entropic* species of irreversibility is only *de facto* rather than *nomological*.

(B) ARE THERE NON-THERMODYNAMIC FOUNDATIONS FOR
THE ANISOTROPY OF TIME?

In a series of notes, published in *Nature* during the years 1956–1958, K. R. Popper² has expounded his thesis of the "untenability of the widespread, though surely not universal, belief that the 'arrow of time' is closely connected with, or dependent upon, the law that disorder (entropy) tends to increase" (II). Specifically, he argues in the first three of his four notes that there exist *irreversible* processes in nature whose irreversibility does not depend on their involvement of an entropy increase. Instead, their irreversibility is *nomologically contingent* in the following sense: the *laws* of nature governing elementary processes do indeed allow the temporal inverses of these irreversible

² K. R. Popper: *Nature*, Vol. CLXXVII (1956), p. 538; Vol. CLXXVIII (1956), p. 382; Vol. CLXXIX (1957), p. 1297; Vol. CLXXXI (1958), p. 402. These four publications will be cited hereafter as "I," "II," "III," and "IV" respectively.

processes, but the latter processes are de facto irreversible, because the *spontaneous* concatenation of the initial conditions requisite for the occurrence of their temporal inverses is well-nigh physically impossible. Noting that "Although the arrow of time is not implied by the fundamental equations [laws governing elementary processes], it nevertheless characterizes most solutions" (I), Popper therefore rejects the claim that "every non-statistical or 'classical' mechanical process is reversible" (IV). In the fourth of his communications, he maintains that the statistical behavior of the entropy of physical systems not only fails to be the *sole* physical basis for the anisotropy of time, as Boltzmann had supposed, but does not qualify *at all* as such a basis.³ For, as we saw earlier, Popper argues that, if it did, the temporal description of fluctuation phenomena would entail absurdities of several kinds.

In response to the first two of Popper's four notes, E. L. Hill and I published a communication⁴ in which we endorsed Popper's contention of the existence of non-entropic, nomologically contingent irreversibility in the form of an existential claim constituting a generalization of Popper's contention.

In view of Popper's criticism (III) of the latter generalization, my aim is to deal with non-entropic irreversibility as follows:

- (1) To appraise Popper's criticism.
- (2) To show that the generalization put forward in the paper by Hill and myself has the important merit of dispensing with the *restriction* on which Popper predicates his affirmation of nomologically contingent irreversibility. This restriction is constituted by the requirement of the *spontaneity* of the concatenation of the initial conditions requisite to the occurrence of the temporal inverses of the thus *conditionally* irreversible processes.
- (3) To assess the import of my appraisal of Popper's claims for Mehlberg's denial of the anisotropy of time.⁵

³ In view of the misleading potentialities of Eddington's metaphor "the arrow of time," which is also employed by Popper, I prefer to substitute the non-metaphorical expression "the anisotropy of time" in my account of Popper's views.

⁴ E. L. Hill and A. Grünbaum: "Irreversible Processes in Physical Theory," *Nature*, Vol. CLXXIX (1957), p. 1296.

⁵ Cf. H. Mehlberg: "Physical Laws and Time's Arrow," *op. cit.*

Independently of O. Costa de Beauregard, who had used the same illustration before him,⁴ Popper (*I*) considers a large surface of water initially at rest into which a stone is dropped, thereby producing an outgoing wave of decreasing amplitude spreading concentrically about the point of the stone's impact. And Popper argues that this process is irreversible in the sense that the "spontaneous" (*IV*) concatenation on all points of a circle of the initial conditions requisite to the occurrence of a corresponding *contracting* wave is physically impossible, a "spontaneous" concatenation being understood to be one which is not brought about by coordinated influences emanating from a common center. Being predicated on the latter spontaneity, this nomologically contingent irreversibility is of a *conditional* kind.

Now, one might object that the attribution of the irreversibility of the outgoing wave motion to non-thermodynamic causal factors is unsound. The grounds would be that the statistical entropy law is *not irrelevant* to this irreversibility, because the diminution in the amplitude of the outgoing wave is due to the superposition of *two* independent effects as follows: First, the requirements of the law of conservation of energy (*first law* of thermodynamics), and second, an entropy increase in an essentially closed system through dissipative viscosity. To be sure, the entropy increase through dissipative viscosity is a *sufficient* condition (in the *statistical* sense of part A. of this chapter!) for the irreversibility of the outgoing wave motion, i.e., for the absence of a corresponding (spontaneously initiated) contracting wave motion. But this fact cannot detract from the soundness of Popper's claim that another, independent *sufficient condition* for the conditional kind of de facto irreversibility affirmed by him is as follows: the nomologically contingent non-existence of the spontaneous occurrence of the coordinated initial conditions requisite for a contracting wave motion. We see that Popper rightly adduces the need for the *coherence* of these initial conditions as his basis for denying the possibility of their *spontaneous* concatenation, i.e., their concatenation without first having been coordinated by an influence emanating from a central

⁴ O. C. de Beauregard: "L'Irréversibilité Quantique, Phénomène Macroscopique," *op. cit.*, p. 402.

source. Says he (III): "Only such conditions can be causally realized as can be organized from one centre . . . causes which are not centrally correlated are causally unrelated, and can co-operate [i.e., produce coherence in the form of *isotropic* contraction of waves to a precise point] only by accident. . . . The probability of such an accident will be zero."

In view of the aforementioned *conditional* character of Popper's nomologically contingent irreversibility, E. L. Hill and I deemed it useful to point out the following:⁷ there does indeed exist an important class of processes in *infinite* space whose *irreversibility* is (1) non-entropic and nomologically contingent, hence being of the kind correctly envisioned by Popper, yet (2) *not conditional* by not being predicated on Popper's proviso of spontaneity. Without presuming to speak for Professor Hill, I can say, for my part, that in making that existential claim I was guided by the following considerations:

1. Popper (II) briefly remarks correctly that the eternal expansion of a very thin gas from a center into a spatially *infinite* universe does not involve an entropy increase, and the de facto irreversibility of this process is therefore *non-entropic*. For the statistical Maxwell-Boltzmann entropy is not even defined for a spatially infinite universe: the quasi-ergodic hypothesis, which provides the essential basis for the probability-metric ingredient in the Maxwell-Boltzmann entropy concept, is presumably *false* for an *infinite* phase-space, since *walls* are required to produce the collisions which are essential to its validity. In the absence of some kind of wall, whose very existence would assure the *finitude* of the system, the rapidly moving particles will soon overtake the slowly moving ones, leaving them ever further behind for all future eternity instead of *mixing* with them in a space-filling manner. Moreover, as we noted earlier, if the number of particles in the infinite universe is only finite, the equilibrium state of maximum entropy cannot be realized, since a finite number of particles cannot be *uniformly* distributed in a phase space of infinitely many cells. Furthermore, if the number of particles is denumerably infinite, the number *W* of micro-

⁷ Cf. E. L. Hill and A. Grünbaum: "Irreversible Processes in Physical Theory," *op. cit.*

scopic complexions in $S = k \log W$ becomes infinite, and no entropy increase or decrease is defined. Corresponding remarks apply to the case of the Bose-Einstein entropy of a gas of photons (bosons) of various frequencies (energies) which is not confined by an enclosure but has free play in an infinite space.⁸

2. Though *allowed* by the *laws* of mechanics, there seem to exist no "implosions" at all which would qualify as the temporal inverses of eternally progressing "explosions" of very thin gases from a center into infinite space. In the light of this fact, one can assert the de facto irreversibility of an *eternal* "explosion" *unconditionally*, i.e., *without* Popper's restrictive proviso of *spontaneity* in regard to the production of the *coherent* initial conditions requisite for its inverse. For in an infinite space, there is even no possibility at all of a *non-spontaneous* production of the coherent "initial" conditions for an implosion having the following properties: the gas particles converge to a point after having been moving through infinite space for all past eternity in a manner constituting the temporal inverse of the expansion of a very thin gas from a point for all future eternity. There can be no question of a *non-spontaneous* realization of the "initial" conditions required for the latter kind of implosion, since such a realization would involve a self-contradictory condition akin to that in Kant's fallacious First Antinomy: the requirement that a process which has been going on for all infinite past time must have had a finite beginning (production by past *initial* conditions) after all.

By contrast, in a spatially *finite* system it is indeed possible to produce *non-spontaneously* the initial conditions for contracting waves and for implosions of gas particles which converge to a point. Thus, assuming negligible viscosity, there are expanding water waves in *finite* systems of which the temporal inverses could be produced *non-spontaneously* by dropping a very large circular object onto the water surface so that all parts of the circular object strike the water surface simultaneously. And hence there are conditions under which contracting waves do exist in finite systems. But there is no need whatever for Popper's

⁸ For a treatment of classical entropic aspects of light propagation within a *finite* system, cf. A. Landé: "Optik und Thermodynamik," *Handbuch der Physik* (Berlin: J. Springer; 1928), Vol. XX, esp. pp. 471-79.

spontaneity proviso to assert the de facto irreversibility of the eternal expansion of a *spherical light wave* from a center through infinite space! If space is infinite, the existence of the latter process of expansion is assured by the facts of observation in conjunction with electromagnetic theory; but despite the fact that the *laws* for a homogeneous and isotropic medium *allow* the *inverse* process no less than the actual one,⁹ we never encounter the inverse process of spherical waves closing in isotropically to a sharp point of extinction.

In view of the decisive role of the infinitude or "openness" of a physical system (the universe)—as opposed to the finitude of closed systems—in rendering Popper's spontaneity proviso dispensable, Hill and I made the following existential claim concerning processes in "open" (infinite) systems whose irreversibility is *non-entropic* and de facto:

In classical mechanics the closed systems have quasi-periodic orbits, whereas the open systems have at least some aperiodic orbits which extend to infinity. . . . there exists a fundamental distinction between the two kinds of system in the following sense. In open systems there always exists a class of allowed elementary processes the inverses of which are unacceptable on physical grounds by requiring a *deus ex machina* for their production. For example, in an open universe, matter or radiation can travel away indefinitely from the "finite" region of space, and so be permanently lost. The inverse process would require matter or radiant energy coming from "infinity," and so would involve a process which is not realizable by physical sources. Einstein's example of an outgoing light wave and Popper's analogous case of a water wave are special finite illustrations of this principle.¹

It will be noted that Hill and I spoke of there being "at least some aperiodic orbits which extend to infinity" in the classical mechanics of open systems and that we were careful not to assert that every such allowed process extending to infinity is a

⁹ Cf. G. J. Whitrow: *The Natural Philosophy of Time* (London: Thomas Nelson & Sons Ltd.; 1961), pp. 8–10 and 269; also E. Zilsel: "Über die Asymmetrie der Kausalität und die Einsinnigkeit der Zeit," *Naturwissenschaften*, Vol. XV (1927), p. 283.

¹ E. L. Hill and A. Grünbaum: "Irreversible Processes in Physical Theory," *op. cit.*

de facto irreversible one. Instead, we affirmed the existence of a de facto irreversibility which is not predicated on Popper's spontaneity proviso by saying: "there always exists a class of allowed elementary processes" that are thus de facto irreversible. And, for my part, I conceived of this claim as constituting an extension of Popper's recognition of the essential role of coherence in de facto irreversibility to processes of the following kind: processes whose de facto irreversibility is not conditional on Popper's finitist requirement of spontaneity, because these processes extend to "infinity" in open systems and would hence have inverses in which matter or energy would have to come from "infinity" *coherently* so as to converge upon a point.

I was therefore quite puzzled to find that the communication by Hill and myself prompted the following dissent by Popper (III):

In this connexion, I must express some doubt as to whether the principle proposed by Profs. Hill and Grünbaum is adequate. In formulating their principle, they operate with two ideas: that of the "openness" of a system, and that of a *deus ex machina*. Both seem to be insufficient. For a system consisting of a sun, and a comet coming from infinity and describing a hyperbolic path around the sun, seems to me to satisfy all the criteria stated by them. The system is open; and the reversion of the comet on its track would require a *deus ex machina* for its realization: it would "require matter . . . coming from 'infinity.'" Nevertheless, this is an example of just that kind of process which, I take it, we all wish to describe as completely reversible.

Popper's proposed counter-example of the comet coming from "infinity" into the solar system seems to me to fail for the following reasons: First, neither the actual motion of the comet nor its inverse involve any *coherence*, a feature which I, for my part, had conceived to be essential to the obtaining of non-entropic de facto irreversibility in open systems. In my own view, the fact that a particle or photon *came from* "infinity" in the course of an infinite past does not per se require a *deus ex machina*, any more than does their *going to* "infinity" in the course of an infinite future: in this context, I regard as innocuous the asymmetry that a particle which has already come from infinity can

be said to have traversed an infinite space *by now*, whereas a particle now embarking on an infinite future journey will only have traversed a *finite* distance at *any one* time in the future. It is a *coherent* "implosion" from infinity that I believe to require a *deus ex machina*, i.e., to be de facto non-existent, while coherent "explosions" actually do exist. Second, even ignoring that the motion of Popper's comet does not involve coherence, the issue is not, as he seems to think, whether it would require a *deus ex machina* to realize the reversal of any given actual comet in its track; rather the issue is whether no *deus ex machina* would be needed to realize the actual comet motion while a *deus ex machina* would have been needed to have another comet execute *instead* a motion inverse to the first one. The answer to this question is an emphatic "no": unlike the case of outgoing and contracting waves (explosions and implosions), the two comet motions, which are temporal inverses of each other, are on a par with respect to the role of a *deus ex machina* in their realization. And even the reversal of the motion of an actual comet at a suitable point in its orbit might in fact be effected by an elastic collision with an oppositely moving other comet of equal mass and hence would not involve, as Popper would have it (III), "a *deus ex machina* who is something like a gigantic tennis player."

It seems to me, therefore, that far from being vulnerable to Popper's proposed counter-example, the existential claim by Hill and myself is fully as viable as Popper's, while having the further merit of achieving generality through freedom from Popper's spontaneity proviso. I therefore cannot see any justification at all for the following two assertions by H. Mehlberg: First, Mehlberg states incorrectly that Hill and I have claimed de facto irreversibility for "the class of all conceivable physical processes provided that the latter meet the mild requirement of happening in an 'open' physical system," and second, Mehlberg asserts that "Popper has shown the untenability of the Hill-Grünbaum criterion by constructing an effective counter example which illustrates the impossibility of their sweeping generalization of his original criterion."²

² H. Mehlberg: "Physical Laws and Time's Arrow," *op. cit.*, p. 128.

Mehlberg's critical estimate of Popper's own affirmation of non-entropic de facto irreversibility likewise seems to me to be unconvincing in important respects. After asking whether the irreversibility asserted by Popper is "lawlike or factlike"—a question to which the answer is: "avowedly factlike"—Mehlberg³ concludes that Popper's temporal asymmetry "seems to be rather interpretable as a local, factlike particularity of the terrestrial surface than as a universal, lawlike feature . . . which may be expected to materialize always and everywhere." There are two points in Mehlberg's conclusion which invite comment: First, the significance which he attaches to the circumstance that the irreversibility of certain classes of processes is de facto or factlike rather than nomological or lawlike, when he assesses the bearing of that irreversibility on the issue of anisotropy *vs.* isotropy of time, and second, the striking contrast between the *epistemological parsimony* of his characterization of Popper's irreversibility as a "local . . . particularity of the terrestrial surface" and the *inductive boldness* of Mehlberg's willingness to do the following: affirm a cosmically pervasive nomological *isotropy* of time on the basis of attributing *cosmic* relevance, both spatially and temporally, to the fundamental *time-symmetric* laws which have been confirmed in modern man's limited sample of the universe.

As to the first of these two points in Mehlberg's denial of the anisotropy of time, I note preliminarily that human hopes for an eternal biological life are no less surely frustrated if all men are indeed de facto mortal, i.e., mortal on the strength of "boundary conditions" which do obtain permanently, than if man's mortality were assured by some law. Moreover, we saw in Chapter Seven that properties of time other than anisotropy depend on boundary conditions and not on the laws alone: the topological openness as opposed to closedness of time is a matter of the boundary conditions, if the laws are deterministic, as is the particular kind of open time that may obtain (finite, half infinite or infinite in both directions). By the same token, I see no escape from the conclusion that *if* de facto irreversibility does actually obtain everywhere and forever, such irreversibility confers anisotropy on time. And this anisotropy prevails not one iota less than

³ *Ibid.*, p. 126.

it would, if its existence were guaranteed by temporally *asymmetrical* fundamental *laws* of cosmic scope.

For what is decisive for the anisotropy of time is not whether the non-existence of the temporal inverses of certain processes is factlike or lawlike; instead, what is relevant for temporal anisotropy is whether the required inverses do actually ever occur or not, whatever the reason. It is of considerable independent interest, of course, that such irreversibility as obtains in nature is *de facto* rather than nomological, *if* the distinction between a law (nomic regularity) and a non-nomic regularity arising from boundary conditions can always be drawn in a conceptually clear way. But, in my view, when evaluating the evidence for the anisotropy of time, Mehlberg commits the following error of misplaced emphasis: he wrongly discounts *de facto* irreversibility vis-à-vis nomological irreversibility by failing to show that our warrant for a cosmic extrapolation of time-symmetric laws is actually greater than for a corresponding extrapolation of the factlike conditions making for observed *de facto* irreversibility. For on what grounds can it be maintained that the ubiquitous and permanent existence of the *de facto* probabilities of "boundary conditions" on which Popper rests his affirmation of temporal anisotropy is less well confirmed than those laws on whose time-symmetry Mehlberg is willing to base his *denial* of the anisotropy of time? In particular, one wonders how Mehlberg could inductively justify his contention that we are only confirming a "particularity of the terrestrial surface," when we find with Popper (III) that:

Only such conditions can be causally realized as can be organized from one centre. . . . causes which are not centrally correlated are causally unrelated, and can co-operate [i.e., produce coherence in the form of *isotropic* contraction of waves to a precise point] only by accident . . . The probability of such an accident will be zero.

If this finding cannot be presumed to hold on all planet-like bodies in the universe, for example, then why are we entitled to assume with Mehlberg that time-symmetric laws of mechanics, for example, are exemplified by the motions of binary stars throughout the universe? Since I see no valid grounds for Mehl-

berg's double standard of inductive credibility of pervasiveness as between laws and factlike regularities, I consider his negative estimate of Popper's non-entropic *de facto* anisotropy of time as quite unfounded.

Mehlberg's misplaced emphasis on the significance of lawlike *vis-à-vis de facto* irreversibility likewise seems to me to vitiate the following account which he gives of the import of *de facto* irreversibility in optics, a species of irreversibility which he admits to be of cosmic scope. He writes:

A less speculative example of cosmological irreversibility is provided by the propagation of light *in vacuo*, which several authors have discussed from this point of view. . . . In accordance with Maxwell's theory of light conceived as an electromagnetic phenomenon, they point out that light emitted by a pointlike source, or converging towards a point, can spread on concentric spherical surfaces which either expand or contract monotonically. Yet, independently of Maxwell's theory, the incidence of expanding optical spheres is known to exceed by far the incidence of shrinking spheres. The reason for this statistical superiority of expanding optical spheres is simply the fact that pointlike light-emitting atoms are much more numerous than perfectly spherical, opaque surfaces capable of generating shrinking optical spheres, mainly by the process of reflection. If true, this ratio of the incidences of both types of light waves would provide a cosmological clue to a pervasive irreversibility of a particular class of optical processes.

The bearing of this optical irreversibility upon time's arrow was often discussed. A long time before the asymmetry of expanding and contracting light waves was promoted to the rank of time's arrow, Einstein⁴ pointed out that the asymmetry of these two types of optical propagation holds only on the undulatory theory of light. Once light is identified instead with a swarm of photons, the asymmetry vanishes. This conclusion holds at least for a spatially finite universe or for optical phenomena confined to a finite spatial region.

Once more, however, the decisive point seems to be that the asymmetry between the two types of light waves depends on factual, initial conditions which prevail in a given momentary

⁴ "A. Einstein: Über die Entwicklung unserer Anschauungen über die Konstitution und das Wesen der Strahlung," *Physikalische Zeitschrift*, Vol. X (1910), pp. 817-28."

cross section of cosmic history or at the "boundaries" of a finite or infinite universe rather than on nomological considerations concerning this history: any other ratio of the incidences of expanding and shrinking light waves would also be in keeping with the relevant laws of nature contained in Maxwell's theory of electromagnetic phenomena. Of course, the aforementioned non-nomological conditions, responsible for the factual ratio of these incidences, are not "local" either, since the whole world is involved—they belong to cosmology. These conditions are nevertheless factlike rather than lawlike, as a comparison with the pertinent laws which can be derived from Maxwell's theory clearly shows.⁵

Contrary to Mehlberg, the decisive point appears not to be that "the asymmetry between the two types of light waves depends on factual, initial conditions . . . rather than on nomological considerations." He also asserts that "at least for a spatially finite universe or for optical phenomena confined to a finite spatial region" the corpuscularity of the photon, as conceived by Einstein, invalidates the optical asymmetry which obtains on the undulatory theory of light. I believe, however, that this claim should be amended as follows: "the optical asymmetry vanishes, *if at all, only* in a *finite* space." For suppose that one assumes with Einstein that the elementary radiation process is one in which a single emitting particle transfers its energy to only a single absorbing particle. In that case, the fantastically complicated coherence needed for the formation of a continuous contracting undulatory spherical shell of light is no longer required. Instead, there is then a need for the less complicated coherence among emitting particles located at the walls of a finite system and emitting converging photons. But, as Hill and I pointed out,⁶ the de facto irreversibility of the spatially symmetrical eternal propagation of a pulse of light from a point source into *infinite* space does not depend on whether the light pulse is undulatory instead of being constituted by a swarm of photons.

Neither does *this* irreversibility depend on the acceptance of

⁵ H. Mehlberg: "Physical Laws and Time's Arrow," *op. cit.*, pp. 123–24.

⁶ E. L. Hill and A. Grünbaum: "Irreversible Processes in Physical Theory," *op. cit.*

the steady state theory of cosmology which, in the words of T. Gold, offers the following explanation for the fact that the universe is a non-reflecting sink for radiation:

It is this facility of the universe to soak up any amount of radiation that makes it different from any closed box, and it is just this that enables it to define the arrow of time in any system that is in contact with this sink. But why is it that the universe is a non-reflecting sink for radiation? Different explanations are offered for this in the various cosmological theories and in some schemes, indeed, this would only be a temporary property.⁷ In the steady state universe it is entirely attributed to the state of expansion. The red shift operates to diminish the contribution to the radiation field of distant matter; even though the density does not diminish at great distances, the sky is dark because in most directions the material on a line of sight is receding very fast. . . .⁸

What Gold appears to have in mind here is that due to a very substantial Doppler shift of the radiation emitted by the receding galaxies, the frequency ν becomes very low or goes to zero, and since the energy of that radiation is given by $E = h\nu$, very little if any is received by us converging from these sources. And he goes on to say:

This photon expansion going on around most material is the most striking type of asymmetry, and it appears to give rise to all other time asymmetries that are in evidence. The preferential divergence, rather than convergence, of the world lines of a system ceases when that system has been isolated in a box which prevents the expansion of the photons out into space. Time's arrow is then lost.

We see that Gold's account includes an appreciation of the decisive role played by the infinitude of the space in rendering irreversible the radiation spreading from a point. To be sure, he does emphasize that the Doppler shift due to the expansion makes for the darkness of the sky at night, which would other-

⁷ Presumably Gold is referring here to models of spatially closed or finite universes.

⁸ T. Gold: "The Arrow of Time," in *La Structure et L'Évolution de l'Univers*, Proceedings of the 11th Solvay Congress (Brussels: R. Stoops, 1958), pp. 86-87.

wise be lit up by strong radiation. But the crucial point is the following: even if the energy of radiation from receding galaxies were not drastically attenuated by the Doppler shift, such radiation would still not be the inverse of a process in which a pulse of photons from a point source forever spreads symmetrically into infinite space from that point source. The non-existent inverse of the latter process of outgoing radiation would be a contracting configuration of photons that has been coming from "infinity," i.e., from no sources at all, and has been converging on a point for all infinite past time.

We see again that the complete time-symmetry of the basic laws like those of dynamics or electromagnetism is entirely compatible with the existence of contingent irreversibility. In the concise and apt words of Penrose and Percival, the reason for this compatibility is that "Dynamics relates the states of a system at two different times, but it puts no restriction whatever on the state at any one time, nor on the probability distribution at any one time."⁹

The anisotropy arising from the non-entropic de facto irreversibility which we have been considering is (1) more pervasive temporally than the merely statistical anisotropy guaranteed by the thermodynamics of branch systems, and (2) more ubiquitous spatially than any exclusively large-scale anisotropy of time, such as would be guaranteed, for example, by the universe's being a monotonically expanding spherical 3-space which had a singular state without temporal predecessor in the finite past. For the non-entropic de facto irreversibility which we have been considering assures uniform temporal anisotropy for local intervals in the time continuum no less than for that continuum in the large. And, unlike such anisotropy as would arise from the monotone expansion of a spherical 3-space type of universe, the anisotropy guaranteed by Popper's non-entropic de facto irreversibility is manifest within the small spatial regions accessible to our daily experience.

The scope of the preceding discussion could be enlarged to include more detailed allowance for the results of further physical theories such as the steady-state cosmology. With respect to

⁹ O. Penrose and I. C. Percival: "The Direction of Time," *Proceedings of the Physical Society*, Vol. LXXIX (1962), p. 606.

that wider class of physical theories, the connection between the thermodynamic and non-thermodynamic kinds of irreversibility may well turn out to be deeper than I have claimed. Specifically, the connection may go beyond the fact that both kinds of irreversibility are due to boundary conditions rather than to laws and fail to obtain in the case of systems which are both spatially finite and permanently closed. Thus, Gold inseparably links all temporal asymmetries which make for the anisotropy of time with the time-asymmetry between the divergence and convergence of the geodesics associated with fundamental observers (expansion of the universe). For he relates the temporal asymmetry in all statistical processes to the tendency for radiation to diverge with positive time, a tendency which he relates, in turn, to the expansion of the universe. In this way, Gold links the thermodynamic and radiational time-asymmetries to a cosmological one.

Since *non-entropic* de facto irreversibility suffices to confer anisotropy on time, the statistically time-asymmetric entropic behavior of branch systems is not a necessary condition for the anisotropy of time. Accordingly, if one says that the latter entropic behavior statistically (and intrinsically) "defines" the relation "later than" for physical, as distinct from psychological (common sense) time, the term "defines" must be construed in the weak sense of "constitutes an empirical indicator." But it is important to realize that the latter weak construal of the entropic "definition" of "later than" is necessitated not by the merely statistical character of the thermodynamic anisotropy of time but by the existence of *non-entropic* irreversibility alongside entropic statistical irreversibility. This important consideration seems to have been overlooked by Carnap in setting forth a criticism of the entropic "definition" of "later than." Carnap writes:

Reichenbach's definition, which is also accepted by Grünbaum,¹ appears to me very problematic. Reichenbach criticizes Boltzmann's definition by pointing out that the correlation be-

¹ My explicit statement of the differences between the "definitions" given by Reichenbach and myself respectively was not available to Carnap when he wrote the comment quoted here. But the reader will recall from the first part of this chapter that despite the obvious and substantial indebtedness of my entropic "definition" of "later than" to that of Reichenbach, there are important respects in which my "definition" does depart from his.

tween the direction of time and the increase in entropy holds, not universally, but only with probability. I agree. But it seems to me that an analogous objection holds for Reichenbach's definition.²

In an unpublished amplification of the latter statement, which I now quote here with his kind permission, Professor Carnap wrote as follows:

If we understand "earlier" in the customary sense of the term in physics, then for a single system the statement "If the entropy at the time point A is considerably lower than at the time point B, then A is earlier than B" holds not universally, but only with probability. But then it seems clear that the same statement with respect to a majority of the branch systems instead of a single branch system, holds likewise only with probability, although under certain conditions the probability may be overwhelmingly great. If the latter is the case, then the increase in entropy may certainly be taken as the basis for an inductive *inference* to the relation E [i.e., "earlier than"]. But it seems very doubtful whether it is legitimate to take a statistical correlation, however high it may be, as the basis for a theoretical *definition*. In other words, if a relation E' is defined in this way, then there are cases, in which E and E' do not agree, where E is understood in the customary sense.

It is quite correct, of course, that my entropic "definition" of "later than" is only statistical: this "definition" presupposes that in most space ensembles of branch systems, the vast majority of the members of the ensemble will increase their entropy in one of the two opposite time directions and decrease it in the other. And these two time directions had already been distinguished from one another to the extent that we utilized the relations of temporal o-betweenness to impose a real number coordinatization extrinsically without assuming at the outset that the entropy statistics of branch systems would turn out to be temporally asymmetric. Hence it is avowedly true that the direction of entropy increase of the majority of branch systems is not the

² R. Carnap: "Adolf Grünbaum on The Philosophy of Space and Time," in: P. A. Schilpp (ed.) *The Philosophy of Rudolf Carnap* (LaSalle: Open Court Publishing Co.; 1963), p. 954. The essentials of my entropic "definition" of "later than" were set forth in my contribution to the latter Schilpp volume entitled "Carnap's Views on the Foundations of Geometry," pp. 599-684.

same for all cosmic epochs which contain branch systems satisfying initial conditions of "randomness" but is the same for only the vast majority of such cosmic epochs. Indeed this fact prompted my speaking of the "statistical" anisotropy of time in this context. I deny altogether, however, Carnap's contention that, on the strength of being statistical, my entropic (and hence intrinsic) "definition" of "later than" (or correspondingly of "earlier than") is beset by the following difficulty charged against it by Carnap: (1) "there are cases, in which E [i.e., the customary "earlier than" of physics as used, for example, to plot the entropy of a permanently closed system against time] and E' [i.e., the entropically "defined" relation of "earlier than"] do not agree," and that therefore (2) "it seems very doubtful whether it is legitimate to take a statistical correlation, however high it might be, as the basis for a theoretical *definition*." My denial of the vulnerability of my "definition" to this criticism of Carnap's is based on the fact that my "definition" of "later than" avowedly utilized the direction of entropy increase for a *typical representative* of the majority of cosmic epochs, so that the ascription of earlier-later relations to states belonging to entropically *atypical* cosmic epochs will be dictated by the fact that the *latter* epochs sustain relations of temporal o-betweenness to the entropically typical epochs. That there is no difficulty here of the kind alleged by Carnap is also apparent from my earlier account of the temporal description of fluctuation phenomena on the basis of my entropic "definition" of "later than": branch systems exhibiting cosmically "occasional" entropy *decreases* with *positive* time can be described as such, since these decreases are temporally *counter-directed* with respect to the entropy increase of the majority of branch systems.

Hence the statistical character of my "definition" of "later than" does not disqualify it. And the valid reason for the rejection by Reichenbach and myself of Boltzmann's attempt at a statistical definition was not that Boltzmann's definition was statistical. Carnap seems to have overlooked that instead, our reason for rejecting Boltzmann's statistics were *completely time-symmetric*, these statistics being those of the long-term entropic behavior of a single, permanently closed system.

Chapter 9

THE ASYMMETRY OF RETRODICTABILITY AND PREDICTABILITY, THE COMPOSSIBILITY OF EXPLANATION OF THE PAST AND PREDICTION OF THE FUTURE, AND MECHANISM VS. TELEOLOGY

The temporally asymmetric character of the entropy statistics of branch systems has a number of important consequences which were not dealt with in Chapter Eight and to which we must now turn our attention. In particular, our conclusions regarding the entropy statistics of branch systems can now be used to elucidate (1) the conditions under which retrodiction of the past is feasible while prediction of the future is not,¹ (2) the relation of psychological time to physical time, (3) the consequence which the feasibility of retrodictability *without* corresponding predictability has for the compossibility of explainability of the past and the corresponding predictability of the future, and (4) the merits of the controversy between philosophical mechanism and teleology.

(A) CONDITIONS OF RETRODICTABILITY AND NON-PREDICTABILITY.

Suppose we encounter a beach whose sand forms a smooth surface except for one place where it is in the shape of a human

¹ The reader will recall that conditions under which the *inverse* asymmetry obtains were discussed in Chapter Eight.

footprint. We know from our previous considerations with high probability that instead of having evolved *isolatedly* from a prior state of uniform smoothness into its present uneven configuration according to the statistical entropy principle for a permanently closed system, the beach was an *open* system in *interaction* with a stroller. And we are aware furthermore that if there is some quasi-closed wider system containing the beach and the stroller, as there often is, the beach achieved its ordered low entropy state of bearing the imprint or interaction-indicator at the expense of an at least compensatory entropy increase in that wider system comprising the stroller: the stroller increased the entropy of the wider system by scattering his energy reserves in making the footprint.

We see that the sandy footprint shape is a genuine indicator and not a randomly achieved form resulting from the unperturbed chance concatenations of the grains of sand. The imprint thus contains information in the sense of being a veridical indicator of an interaction. Now, in all probability the entropy of the imprint-bearing beach-system increases after the interaction with the stroller through the smoothing action of the wind. And this entropy increase is parallel, in all probability, to the direction of entropy increase of the majority of branch systems. Moreover, we saw that the production of the indicator by the interaction is likely to have involved an entropy increase in some wider system of which the indicator is a part. Hence, *in all probability the states of the interacting systems which do contain the indicators of the interaction are the relatively higher entropy states of the majority of branch systems, as compared to the interaction state. Hence the indicator states are the relatively later states as compared to the states of interaction which they attest.* And by being both *later* and indicators, these states have *retrodictive* significance, thereby being traces, records or "memories." And due to the high degree of retrodictive univocity of the low entropy states constituting the indicators, the latter are veridical to a high degree of *specificity*.

Confining our attention for the present to indicators whose *production* requires only the occurrence of the interaction which they attest, we therefore obtain the following conclusion. Apart from some kinds of *advance*-indicators requiring very special

conditions for their production and constituting exceptions, it is the case that *with overwhelming probability, low entropy indicator-states can exist in systems whose interactions they attest only after and not before these interactions.*² If this conclusion is true (assuming that there are either no cases or not enough cases of bona fide precognition to disconfirm it), then, of course, it is not an a priori truth. And it would be very shallow indeed to seek to construe it as a trivial a priori truth in the following way: calling the indicator states "traces," "records," or "memories" and noting that it then becomes tautological to assert that traces and the like have only retrodictive and no predictive significance. But this transparent verbal gambit cannot make it true a priori that—apart from the exceptions to be dealt with below—interacting systems bear indicators attesting veridically only their earlier and not their later interactions with outside agencies.

Hence, the exceptions apart, we arrive at the fundamental asymmetry of recordability: *reliable indicators in interacting systems permit only retrodictive inferences concerning the interactions for which they vouch but no predictive inferences pertaining to corresponding later interactions.*

And the logical schema of these inductive inferences is roughly as follows: the premises assert (1) the presence of a certain relatively low entropy state in the system, and (2) the quasi-universal statistical law stating that most low entropy states are interaction-indicators *and* were *preceded* by the interactions for which they vouch. The conclusion from these premises is then the inductive retrodictive one that there was an earlier interaction of a certain kind.

As already mentioned, our affirmation of the temporal asymmetry of recordability of interactions must be qualified by dealing with two exceptional cases, the first of which is the pre-recordability of those interactions which are veridically predicted by human beings (or computers). For any event which

² Two of the exceptions, which we shall discuss in some detail below, are constituted by the following two classes of advance indicators: First, veridical predictions made and stored (recorded) by human (or other sentient, theory-using) beings, and physically registered, bona fide advance indicators produced by computers, and second, advance indicators (e.g., sudden barometric drops) which are produced by the very cause (pressure change) that also produces the future interaction (storm) indicated by them.

could be predicted by a scientist could also be "pre-recorded" by that scientist in various forms such as a written entry on paper asserting its occurrence at a certain later time, an advance drawing, or even an advance photograph based on the pre-drawing. By the same token, artifacts like computers can pre-record events which they can predict. A comparison between the written, drawn, or photographic pre-record (i.e., recorded prediction) of, say, the crash of a plane into a house and its post-record in the form of a caved-in house, and a like comparison of the corresponding pre- and post-records of the interaction of a foot with a beach will now enable us to formulate the essential differences in the conditions requisite to the respective production of pre-records and post-records as well as the usual differences in make-up between them.

The production of at least one retrodictive indicator or post-record of an interaction such as the plane's crash into the house requires only the occurrence of that interaction (as well as a moderate degree of durability of the record). The retrodictive indicator states in the system which interacted with an outside agency must, of course, be distinguished from the *epistemic use* which human beings may make of these physical indicator states. And our assertion of the sufficiency of the interaction for the production of a post-record allows, of course, that the *interpretation* of actual post-records by humans as bona fide documents of the past requires their use of theory and not just the occurrence of the interaction. In contrast to the sufficiency of an interaction itself for its (at least short-lived) post-recordability, no such sufficiency obtains in the case of the pre-recordability of an interaction: save for an overwhelmingly improbable freak occurrence, the production of even a single pre-record of the coupling of a system with an agency external to it requires, as a necessary condition, *either* (a) the use of an appropriate theory by symbol-using entities (humans, computers) having suitable information, *or* (b) the pre-record's being a partial effect of a cause that also produces the pre-recorded interaction, as in the barometric case to be dealt with below. And in contexts in which (a) is a necessary condition, we find the following: since pre-records are, by definition, veridical, this necessary condition cannot *generally* also be sufficient, unless the predictive theory employed is deter-

ministic *and* the information available to the theory-using organism pertains to a closed system.

In addition to differing in regard to the conditions of their production, pre-records generally differ from post-records in the following further respect: unless the pre-record prepared by a human being (or computer) *happens* to be part of the interacting system to which it pertains, the pre-record will not be contained in states of the interacting system which it concerns but will be in some other system. Thus, a pre-record of the crash of a plane into a house in a heavy fog would generally *not* be a part of either the house or the plane, although it can happen to be. But in the case of *post*-recording, there will always be at least one post-record, however short-lived, in the interacting system itself to which that post-record pertains.

Our earlier example of the footprint on the beach will serve to illustrate more fully the asymmetry between the requirements for the production of a pre-record and of a post-record. The pre-recording of a *later* incursion of the beach by a stroller would require extensive information about the motivations and habits of people not now at the beach and also knowledge of the accessibility of the beach to prospective strollers. This is tantamount to knowledge of a large system which is *closed*, so that all relevant agencies can safely be presumed to have been included in it. For otherwise, we would be unable to guarantee, for example, that the *future* stroller will not be stopped enroute to the beach by some agency not included in the system, an eventuality whose occurrence would deprive our pre-record of its referent, thereby destroying its status as a veridical indicator. In short, in the case of the footprint, which is a post-record and not a pre-record of the interaction of a human foot with the beach, the interaction itself is *sufficient* for its post-recording (though not for the extended *durability* of the record once it exists) but *not* for its pre-recording and prediction. Since a future interaction of a potentially open system like the beach is not itself sufficient for its pre-recordability, open systems like beaches therefore do not themselves exhibit pre-records of their own future interactions. Instead—apart from the second species of pre-recordability to be considered presently—pre-recordability of interactions of potentially open systems requires the mediation

of symbol and theory-using organisms or the operation of appropriate artifacts like computers. And such pre-recordability can obtain successfully only if the theory available to the pre-recording organism is deterministic and sufficiently comprehensive to include all the relevant laws and boundary conditions governing the pertinent closed systems.

The second species of exceptions to the asymmetry of recordability is exemplified by the fact that a sudden drop in the pressure reading of a barometer can be an advance-indicator or "pre-record" of a subsequent storm. To be sure, it is the immediately *prior* pressure change in the spatial vicinity of the barometer and only that particular prior change (i.e., the *past* interaction through pressure) which is recorded numerically by a given drop in the barometric reading, and not the pressure change that *will* exist at that same place at a *later* time. To make the predictions required for a *pre*-recording of the pressure changes which will exist at a given space point at later times (i.e., of the corresponding future interactions), comprehensive meteorological data pertaining to a large region would be essential. But it is possible in this case to base a rather reliable prediction of a future storm on the present sudden barometric drop. The latter drop, however, is, in fact, a bona fide advance indicator only because it is a partial effect of the very comprehensive cause which also produces (assures) the storm. Thus, it is the fulfillment of the *necessary condition* of having a causal ancestry that overlaps with that of the storm which is needed to confer the status of an advance indicator on the barometric drop. In contrast to the situation prevailing in the case of *post*-recordability, the existence of this necessary condition makes for the fact that the future occurrence of a storm is *not itself sufficient* for the existence of an advance indicator of that storm in the form of a sudden barometric drop at an earlier time.

An analogous account can be given of the following case, which Mr. F. Brian Skyrms has suggested to me for consideration: situations in which *human intentions* are highly reliable advance indicators of the events envisaged by these intentions. Thus, the desire for a glass of beer, coupled with the supposed presence of the conditions under which beer and a glass are obtainable, produces as a partial effect the intent to get it. And, if external conditions permit (the beer is available and acces-

sible), and, furthermore, if the required internal conditions materialize (the person desiring the beer remains able to go and get it), then the intent will issue in the obtaining and drinking of the beer. But in contrast to the situation prevailing in the case of retrodictive indicators (post-records), the future consumption of the beer is not a sufficient condition for the existence of its probabilistic advance indicator in the form of an intention.

The consideration of some alleged counter-examples will serve to complete our statement of the temporal asymmetry of the recordability of interactions. These purported counter-examples are to the effect that there are pre-records not depending for their production on the use of predictive theory by symbol-using organisms or on the pre-record's being a partial effect of a cause that also produces the pre-recorded interaction.

In the first place, it might be argued that there are spontaneous pre-records as exemplified in the following two kinds of scientific contexts: first, in any essentially closed dynamical system such as the solar system, a dynamical state later than one occurring at a time t_0 is a sufficient condition for the occurrence of the state at time t_0 , no less than is a state prior to t_0 ; hence the state at time t_0 can be regarded as a pre-record of the later state no less than it can be deemed a post-record of the earlier one, and second, a certain kind of death—say, the kind of death ensuing from leukemia—may be a sufficient condition for the existence of a pre-record of it in the form of the onset of active leukemia. But these examples violate the conditions on which our denial of spontaneous pre-recordability is predicated in the following essential respect: they involve later states which are not states of interaction with outside agencies entered into by an otherwise closed system, in the manner of our example of the beach.

In the second place, since the thesis of the temporal asymmetry of spontaneous recordability makes cases of bona fide precognition overwhelmingly improbable, it might be said that this thesis and the entropic considerations undergirding it are vulnerable to the discovery of a reasonable number of cases of genuine precognition, a discovery which is claimed by some to have already been made. To this I retort that if the purported occurrence of precognition turns out to become well authenticated, then I am, of course, prepared to envision such alterations

in the body of current orthodox scientific theory as may be required.

The connection between low entropy states and retrodictive information which emerged in our discussion of the asymmetry of recordability throws light on the reason for the failure of Maxwell's well-known sorting demon: in substance, Maxwell's demon cannot violate the second law of thermodynamics, since the entropy decrease produced in the gas is more than balanced by the entropy increase in the mechanism procuring the informational data concerning individual gas molecules which are needed by the demon for making the sorting successful.³

We saw earlier how reliance on entropy enables us to ascertain which one of two causally connected events is *the* cause of the other because it is the earlier of the two. Our present entropic account of the circumstances under which the past can be inferred from the present while the future cannot, as well as our earlier statement in Chapter Eight of the circumstances when only the converse determination is possible enables us to specify the conditions of validity for the following statements by Reichenbach: "Only the totality of all causes permits an inference concerning the future, but the past is inferable from a partial effect alone" and "one can infer the total cause from a partial effect, but one cannot infer the total effect from a partial cause."⁴ A partial effect produced in a system while it is open permits, on entropic grounds, an inference concerning the earlier interaction event which was its cause: even though we do not

³ Cf. L. Brillouin: *Science and Information Theory*, *op. cit.*; E. C. Cherry: "The Communication of Information," *American Scientist*, Vol. XL (1952), p. 640; J. Rothstein: "Information, Measurement and Quantum Mechanics," *Science*, Vol. CXIV (1951), p. 171, and M. S. Watanabe: "Über die Anwendung Thermodynamischer Begriffe auf den Normalzustand des Atomkerns," *Zeitschrift für Physik*, Vol. CXIII (1939), pp. 482-513.

⁴ H. Reichenbach: "Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft," *Berichte der Bayerischen Akademie München, Mathematisch-Naturwissenschaftliche Abteilung* (1925), p. 157, and "Les Fondements Logiques de la Mécanique des Quanta," *op. cit.*, p. 146. Cf. also C. F. von Weizsäcker: "Der Zweite Hauptsatz und der Unterschied von Vergangenheit und Zukunft," *op. cit.* By noting in his later publications (especially in *DT*, *op. cit.*, pp. 157-67) that the temporal asymmetry involved here has an entropic basis, Reichenbach abandoned his earlier view that it provides an *independent* criterion for the anisotropy of time. Thus, he has essentially admitted the validity of H. Bergmann's telling criticisms (*Der Kampf um das Kausalgesetz in der jüngsten Physik*, *op. cit.*, pp. 19-24) of his earlier view.

know the total present effect, we know that the part of it which is an ordered, low entropy state was (most probably) preceded by a still lower entropy state and that the diversity of the interactions associated with such a very low interaction entropy state is relatively small, thereby permitting a rather specific assertion about the past.

Thus, the asymmetry of inferability arises on the macro-level in the absence of knowledge of the microscopic state of the total (closed) system at a given time and is made possible by the relative retrodictive univocity of local low entropy states which result from interactions. We are therefore in possession of the answer to the question posed by J. J. C. Smart when he wrote: "Even on a Laplacian view, then, we still have the puzzling question: 'Why from a limited region of space can we deduce a great deal of the history of the past, whereas to predict similar facts about the future even a superhuman intelligence would have to consider initial conditions over a very wide region of space?'"⁵

(B) THE PHYSICAL BASIS FOR THE ANISOTROPY OF PSYCHOLOGICAL TIME.

We saw that the production of traces, records, or memories is usually accompanied by entropy increases in the branch systems. And therefore it follows that the direction of increase of stored information or "memories" either in inanimate recording-devices or in memory-gathering organisms like man must be the same as the direction of entropy increase in the majority of branch systems. But we noted in our discussion of Bridgman's critique of Eddington in Chapter Eight that the direction of the increase or accumulation of mental traces or memories is the *forward* direction of psychological time. Hence, what is *psychologically* later goes hand-in-hand with what is purely physically later on the basis of the entropic evolution of branch systems. And the anisotropy of psychological time *mirrors* the more fundamental anisotropy of physical time.

By way of concrete example, note that the vocal output of a lecturer increases the entropy of the lecture room and produces

⁵ J. J. C. Smart: "The Temporal Asymmetry of the World," *op. cit.*, p. 81.

memories in his listeners at his physiological expense. At the start of his lecture, the lecturer had concentrated energy, and therefore the entropy of the lecture room was then relatively low. After his lecture much of that energy has been scattered in the form of sound waves, thereby increasing the entropy of the room and also registering memories in his listeners which enable them to tell that it is then later than it was when they first sat down for the lecture.

These considerations show that even if our physiological existence were possible in a state of total equilibrium—which it is not—we would then *not* be able to have the kind of temporal awareness that we do have. If we were unfortunate enough to be surviving while immersed in one vast cosmic equilibrium, it would then be unutterably dull to the point of loss of our psychological time sense.

See Append. § 30

(C) THE BEARING OF RETRODICTABILITY AND NON-PREDICTABILITY ON THE COMPOSSIBILITY OF EXPLAINABILITY AND PREDICTABILITY.

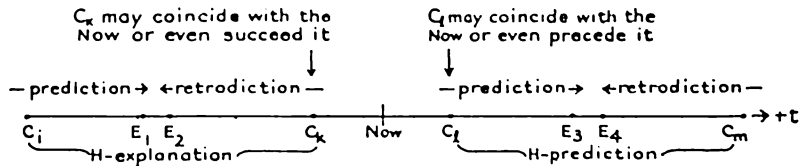
What is the bearing of the existence of the asymmetry between retrodiction and prediction on the following quite distinct question: Is there symmetry between the *explanation* of an event E on the basis of one or more *antecedents* of E, when E belongs to the *past* of the explaining scientist, on the one hand, and, on the other hand, the *prediction* of the same (kind of) E by reference to the same (kind of) *antecedent(s)* of E, when E belongs to the *future* of the scientist making the prediction? In short, what is the relevance of the retrodiction-prediction asymmetry to the thesis of symmetry (or structural equality) of explanation and prediction as put forward by writers such as C. G. Hempel?⁶ Preparatory to answering this question, we shall clarify the temporal relations which are involved and will then represent the results on a diagram.

⁶ I refer here to the original paper of C. G. Hempel and P. Oppenheim: "Studies in the Logic of Explanation," *Philosophy of Science*, Vol. XV (1948), p. 15. For Hempel's most recent statement of his account of scientific explanation, see his "Deductive Nomological vs. Statistical Explanation," in: H. Feigl and G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science* (Minneapolis: University of Minnesota Press; 1962), Vol. III, pp. 98-169.

For Hempel, the particular conditions C_i ($i = 1, 2, \dots, n$) which, in conjunction with the relevant laws, account for the *explanandum*-event E , may be *earlier* than E in *both* explanation and prediction or the C_i may be *later* than E in *both* explanation and prediction. Thus, a case of *prediction* in which the C_i would be *later* than E would be one in astronomy, for example, in which a future E is accounted for by reference to C_i which are still further in the future than E . These assertions hold, since Hempel's criterion for an explanation as opposed to a prediction is that E belong to the scientist's *past* when he offers his account of it, and his criterion for a corresponding prediction is that E belong to the scientist's *future* when it is made.

However, in the retrodiction-prediction antithesis, a *retrodiction* is characterized by the fact that the C_i are *later* than E , while the C_i are *earlier* than E in the kind of *prediction* which is antithetical to retrodiction but *not* identical with Hempelian prediction.

In the accompanying diagram, the i, k, l, m , may each range over the values $1, 2 \dots n$.



If we use the prefix "H" as an abbreviation for "Hempelian," then two consequences are apparent. Firstly, a retrodiction as well as a prediction can be an H-prediction, and a prediction as well as a retrodiction can be an H-explanation. Secondly, being an H-prediction rather than an H-explanation or conversely depends on the transient homocentric "now," but there is no such "now"-dependence in the case of being a retrodiction instead of a prediction, or conversely.

The passage in the Hempel-Oppenheim essay setting forth the symmetry thesis espoused by K. R. Popper and these authors reads as follows:

the same formal analysis, including the four necessary conditions, applies to scientific prediction as well as to explanation. The difference between the two is of a pragmatic character. If E is given, i.e., if we know that the phenomenon described by

E has occurred, and a suitable set of statements $C_1, C_2, \dots, C_k, L_1, L_2, \dots, L_r$ is provided afterwards, we speak of an explanation of the phenomenon in question. If the latter statements are given and E is derived prior to the occurrence of the phenomenon it describes, we speak of a prediction. It may be said, therefore, that an explanation is not fully adequate unless its explanans, if taken account of in time, could have served as a basis for predicting the phenomenon under consideration.⁷ Consequently, whatever will be said in this article concerning the logical characteristics of explanation or prediction will be applicable to either, even if only one of them should be mentioned.⁸

Hempel's thesis of symmetry or structural equality between H-explanation and H-prediction can therefore now be formulated in the following way: Any *prediction* which qualifies logically *and* methodologically as an H-explanation also qualifies as an H-prediction, provided that the scientist is in possession of the information concerning the C_i prior to the occurrence of E, and conversely. And any *retrodiction* which qualifies logically *and* methodologically as an H-explanation also qualifies as an H-prediction, provided that the information concerning the relevant C_i is available at an appropriate time, and conversely.

Before examining critically the diverse objections which have been leveled against Hempel's thesis of symmetry in the recent literature by N. Rescher,⁹ S. F. Barker,¹ N. R. Hanson,² and M. Scriven,³ I wish to make a few remarks concerning my con-

⁷ "The logical similarity of explanation and prediction, and the fact that one is directed towards past occurrences, the other towards future ones, is well expressed in the terms 'postdictability' and 'predictability' used by Reichenbach in [Quantum Mechanics], p. 13."

⁸ C. G. Hempel and P. Oppenheim: "Studies in the Logic of Explanation," *op. cit.*, §3.

⁹ N. Rescher: "On Prediction and Explanation," *British Journal for the Philosophy of Science*, Vol. VIII (1958), p. 281.

¹ S. F. Barker: "The Role of Simplicity in Explanation," in: H. Feigl & G. Maxwell (eds.) *Current Issues in the Philosophy of Science*, *op. cit.*, pp. 265-86 and the Comments on this paper by W. Salmon, P. K. Feyerabend, and R. Rudner with Barker's Rejoinders.

² N. R. Hanson: "On the Symmetry Between Explanation and Prediction," *The Philosophical Review*, Vol. LXVIII (1959), p. 349.

³ M. Scriven: "Explanation and Prediction in Evolutionary Theory," *Science*, Vol. CXXX (1959), pp. 477ff., excerpts from which are reprinted from *Science* by permission, and "Explanations, Predictions and Laws," in: H. Feigl & G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science*, Vol. III, *op. cit.*, pp. 170-230.

strual of both that thesis and of the philosophical task to whose fulfillment it pertains.

I take Hempel's affirmation of symmetry to pertain not to the *assertibility* per se of the *explanandum* but to the either deductive or inductive *inferability* of the *explanandum* from the *explanans*. Popper and Hempel say: To the extent that there is ever explanatory *inferability*, there is also predictive inferability and conversely. They do not claim that every time you are entitled to assert, on some grounds or other, that a certain kind of event did occur in the past, you are also entitled to say that the same kind of event will occur in the future. Being concerned with scientific understanding, Popper and Hempel said that there is temporal symmetry not of assertibility per se but of assertibility *on the strength of the explanans*. The scientific relevance of dealing with predictive arguments rather than mere predictive assertions can hardly be contested by claiming with Scriven that in this context "the crucial point is that, however achieved, a prediction is what it is simply because it is produced in advance of the event it predicts; it is *intrinsically* nothing but a bare description of that event."⁴ For surely a soothsayer's unsupported prophecy that there will not be a third world war is not of scientific significance and ought not to command any scientific interest precisely because of the unreasoned manner of its achievement. Hence a scientifically warranted prediction of an event must be more than a mere pre-assertion of the event. And in any context which is to be scientifically relevant, the following two components can be distinguished in the meaning of the term "H-predict" no less than in the meaning of "H-explain" (or "post-explain"), and similarly for the corresponding nouns:

(1) the mere *assertion* of the *explanandum*, which *may* be based on grounds other than its scientific *explanans*,

(2) the logical *derivation* (deductive or inductive) of the *explanandum* from an *explanans*, the character of the content of the *explanans* remaining unspecified until later on in this section.

My attachment of the prefix "H" to the word "explain" (and to "explanation") and my use of "post-explain" as a synonym of "H-explain" will serve to remind us for the sake of clarity that this usage of "explain" results from a restriction to the past of *one*

⁴ M. Scriven: "Explanations, Predictions and Laws," *op. cit.*, Sec. 3.4.

well-established usage which is *temporally neutral*, viz., "explain" in the sense of providing scientific understanding (or a scientific accounting) of why something did *or* will occur. But, to my mind, the philosophical task before us is not the ascertainment of how the words "explain" and "predict" are used, even assuming that there is enough consistency and precision in their usage to make this lexicographic task feasible. And hence the verdict on the correctness of Hempel's symmetry thesis cannot be made to depend on whether it holds for what is taken to be the actual or ordinary usage of these terms. Instead, in this context I conceive the philosophical task to be both the elucidation and examination of the provision of scientific understanding of an *explanandum* by an *explanans* as encountered in actual scientific theory. Accordingly, Hempel's symmetry thesis, which concerns the inferability of the *explananda* from a given kind of *explanans* and not their assertibility, must be assessed on the basis of a comparison of H-predictive with H-explanatory *arguments* with respect to the measure of scientific understanding afforded by them. Thus, the issue of the adequacy of the symmetry thesis will revolve around whether there is temporal symmetry in regard to the degree of entailment, as it were, characterizing the logical link between the *explanans* and the *explanandum*. Specifically, we shall need to answer both of the following questions: First, would the type of *argument* which yields a prediction of a future *explanandum*-event not furnish precisely the same amount of scientific understanding of a corresponding past event? And second, does an *explanans* explain an *explanandum* referring to a past event any more conclusively than this same kind of *explanans predictively implies* the *explanandum* pertaining to the corresponding future event?

We are now ready to turn to the appraisal of the criticisms of Hempel's symmetry thesis offered by Rescher, Barker, Hanson, and Scriven. In the light of my formulation of Hempel's thesis, it becomes clear that it does *not* assert, as Rescher supposes, that any set of C_i which permit a *predictive* inference also qualify for a corresponding *retrodictive* one, or that the converse is true. As Rescher notes correctly but irrelevantly, whether or not symmetry obtains between prediction and retrodiction in any given domain of empirical science is indeed not a purely logical ques-

tion but depends on the content of the laws pertaining to the domain in question. We see therefore that Hempel was justified in claiming⁵ that Rescher has confused H-explanation with retrodiction. And this confusion is also facilitated by one of Scriven's statements of the symmetry thesis, which reads: "to predict, we need a correlation between present events and future ones—to explain, between present ones and past ones."⁶

In agreement with Scheffler,⁷ Rescher offers a further criticism of the Hempelian assertion of symmetry: "it is inconsistent with scientific custom and usage regarding the concepts of explanation and prediction," for, among other things, "Only true statements are proper objects for explanation, but clearly not so with prediction."⁸ Thus, Rescher notes that we explain only phenomena which are known to have occurred, but we sometimes predict occurrences which do not come to pass. And in support of the latter claim of an "epistemological asymmetry," Rescher points to a large number of cases in which we have "virtually certain knowledge of the past on the basis of traces found in the present" but "merely probable knowledge of the future on the basis of knowledge of the present and/or the past."⁹

The question raised by Rescher's further objection is whether this *epistemological* asymmetry can be held to impugn the Hempelian thesis of symmetry. To deal with this question, it is fundamental to distinguish—as Rescher, Barker, Hanson, and Scriven unfortunately failed to do, much to the detriment of their theses—between the following two sets of ideas: First, an asymmetry between H-explanation and H-prediction both in regard to the grounds on which we claim to know that the *explanandum* is *true* and correlatively in regard to the *degree* of our *confidence* in the supposed truth of the *explanandum*, and second, an asymmetry, if any, between H-explanation and H-prediction with respect to the *logical relation* obtaining between the *explanans*

⁵ C. G. Hempel: "Deductive-Nomological vs. Statistical Explanation," *op. cit.*, Sec. 6.

⁶ M. Scriven: "Explanation and Prediction in Evolutionary Theory," *op. cit.*, p. 479.

⁷ I. Scheffler: "Explanation, Prediction and Abstraction," *British Journal for the Philosophy of Science*, Vol. VII (1957), p. 293.

⁸ N. Rescher: "On Prediction and Explanation," *op. cit.*, p. 282.

⁹ *Ibid.*, p. 284.

and the *explanandum*. For the sake of brevity, we shall refer to the first asymmetry as pertaining to the "assertibility" of the *explanandum* while speaking of the second as an asymmetry in the "inferability" or "why" of the *explanandum*. In the light of this distinction, we shall be able to show that the existence of an epistemological asymmetry in regard to the assertibility of the *explanandum* cannot serve to impugn the Hempelian thesis of symmetry, which pertains to only the why of the *explanandum*.

If understood as pertaining to the *assertibility* per se of the *explanandum*, Rescher's contention of the existence of an epistemological asymmetry is indeed correct. For we saw in the first section of this chapter that there are *highly reliable* retrodictive indicators or records of past interactions but generally no spontaneously produced pre-indicators of corresponding future interactions. And this fact has the important consequence that while we can certify the *assertibility* or truth of an *explanandum* referring to a *past* interaction on the basis of a record without invoking the supposed truth of any (usual) *explanans* thereof, generally no pre-indicator but only the supposed truth of an appropriate *explanans* can be invoked to vouch for the assertibility or truth of the *explanandum* pertaining to a *future* interaction. And since the theory underlying our interpretations of records is confirmed better than are many of the theories used in an *explanans*, there is a very large class of cases in which an epistemological asymmetry or asymmetry of recordability does obtain with respect to the assertibility of the *explanandum*. But this asymmetry of *assertibility* cannot detract from the following *symmetry* affirmed by Popper and Hempel: To the same extent to which an *explanandum* referring to the past can be post-asserted *on the strength of its explanans* in an H-explanation, a corresponding *explanandum* referring to the future can be pre-asserted on the strength of the *same explanans* in an H-prediction. In other words, you can post-assert an *explanandum on the strength of its explanans* no better than you can pre-assert it.

The entire substance of both Barker's objection to Hempelian symmetry and of Hanson's (1959) critique of it is vitiated by the following fact: these authors adduced what they failed to recognize as a temporal asymmetry in the mere assertibility of the

explanandum to claim against Hempel that there is a temporal asymmetry in the *why*. And they did so by citing cases in which they invoke a spurious contrast between the *non-assertibility* of an *explanandum* referring to the future and the *inductive inferability* of the corresponding *explanandum* pertaining to the past. Thus we find that Barker writes: "it can be correct to speak of explanation in many cases where specific prediction is not possible. Thus, for instance, if the patient shows all the symptoms of pneumonia, sickens and dies, I can then explain his death—I know what killed him—but I could not have definitely predicted in advance that he was going to die; for usually pneumonia fails to be fatal."¹ But all that Barker is entitled to here is the following claim, which is wholly compatible with Hempel's symmetry thesis: in many cases such as the pneumonia one, there obtains post-assertibility of the *explanandum* in the presence of the corpse but no corresponding pre-assertibility because of the asymmetry of spontaneous recordability. But this does not, of course, justify the contention that a past death, which did materialize and is reliably known from a record, can be explained by reference to earlier pneumonia any more conclusively than a future death can be inferred predictively on the basis of a present state of pneumonia. For the logical link between the *explanans* affirming a *past* state of pneumonia and the *explanandum* stating the recorded (known) death of a pneumonia patient is precisely the same inductive one as in the following case: the corresponding avowedly probabilistic *predictive* inference (H-prediction) of death on the basis of an *explanans* asserting a patient's *present* affliction with pneumonia.

It would seem that the commission of Barker's error of affirming an asymmetry in the *why* is facilitated by the following question-begging difference between the *explanans* used in his H-explanation of a death from pneumonia and the one used by him in the purportedly corresponding prediction: Barker's H-explanation of the past death employs an *explanans* asserting the onset of pneumonia at a past time as well as the sickening at a later past time, but the further condition of sickening is

¹ S. F. Barker: "The Role of Simplicity in Explanation," *op. cit.*, p. 271.

omitted from the antecedents of his corresponding H-prediction. Hence the spurious asymmetry of conclusiveness between the two cases.

It is now apparent that the valid core of Barker's statement is the commonplace that in the pneumonia case, as in others, post-assertibility of the *explanandum* does obtain even though pre-assertibility does not. And once it is recognized that the only relevant asymmetry which does obtain in cases of the pneumonia type is one of assertibility on grounds other than the usual *explanans*, the philosophical challenge of this asymmetry is to specify the complex reasons for it, as I have endeavored to do above. But no philosophical challenge is posed for Hempel's symmetry thesis.

An analogous confusion between the assertibility asymmetry and one in the *why* invalidates the paper by Hanson which Barker cites in support of his views. Suppose that a certain kind of past measurement yielded a particular Ψ -function which is then used in Schrödinger's equation for the H-explanation of a later past occurrence. And suppose also that the same kind of present measurement again yields the same Ψ -function for a like system and that this function is then used for the H-prediction of a correspondingly later future occurrence, which is of the same type as the past occurrence. It is patent that in quantum mechanics the *logical relation* between *explanans* (the function Ψ_1 and the associated set s_1 of probability distributions at the time t_1) and *explanandum* (the description of a particular micro-event falling within the range of one of the s_1 probability distributions) is no less statistical (inductive) in the case of H-explanation than in the case of H-prediction. And this *symmetry* in the *statistical why* is wholly compatible with the following asymmetry: the reliability of our knowledge *that* a specific kind of micro-event belonging to the range of one of the s_1 probability distributions has occurred in the past has no counterpart in our knowledge of the future occurrence of such an event, because only the results of past measurements (interactions) are available in records.

Hence it was wholly amiss for Hanson to have used the latter asymmetry of recordability as a basis for drawing a *pseudo-contrast* between the quantum mechanical *inferability* of a past

micro-event—this inferability being logically identical with that of a future one—and the lack of *pre-assertibility* of the future occurrence of the micro-event. Says he: “any single quantum phenomenon P . . . can be completely *explained ex post facto*; one can *understand* fully just what kind of event occurred, in terms of the well-established laws of the . . . quantum theory. . . . But it is, of course, the most fundamental feature of these laws that the *prediction* of such a phenomenon P is, as a matter of theoretical principle, quite impossible.”² Hanson overlooks that the asymmetry between pre-assertibility and post-assertibility obtaining in quantum mechanics in no way makes for an asymmetry between H-explanation and H-prediction with respect to the relation of the *explanandum* to its quantum mechanical *explanans*. And the statistical character of quantum mechanics enters only in the following sense: when coupled with the recordability asymmetry of classical physics, it makes for a temporal asymmetry in the assertibility of the *explanandum*.

We see that the statistical character of the quantum mechanical account of micro-phenomena is no less compatible with the symmetry between H-explanation and H-prediction than is the *deterministic* character of Newton’s mechanics. And this result renders untenable what Hanson regards as the upshot of his 1959 paper on the symmetry issue, viz., “that there is a most intimate connection between Hempel’s account of the symmetry between explanation and prediction and the logic of Newton’s *Principia*.”³

It remains to deal in some detail with Scriven’s extensive critique of Hempel’s thesis. Scriven argues that (1) evolutionary explanations and explanations like that of the past occurrence of paresis due to syphilis fail to meet the symmetry requirement by not allowing corresponding predictions, (2) predictions based on mere *indicators* (rather than causes) such as the prediction of a storm from a sudden barometric drop are not matched by corresponding explanations, since indicators are not explanatory though they may serve to predict (or, in other cases, to retrodict). And, according to Scriven, these indicator-based predictions show that the mere inferability of an *explanandum* does not

² N. R. Hanson: “On the Symmetry Between Explanation and Prediction,” *op. cit.*, pp. 353–54.

³ *Ibid.*, p. 357.

guarantee scientific understanding of it, so that symmetry of inferability does not assure symmetry of scientific understanding between explanation and prediction.

I shall now examine several of the paradigm cases adduced by Scriven in support of these contentions.

I. *Evolutionary Theory.*

He cites evolutionary theory with the aim of showing that "Satisfactory explanation of the past is possible even when prediction of the future is impossible."⁴

Evolutionary theory does indeed afford valid examples of the epistemological asymmetry of assertibility. And this for the following two reasons growing out of Section A of the present chapter: First, the ubiquitous role of interactions in evolution brings the recordability asymmetry into play. And that asymmetry enters not only into the assertibility of the *explanandum*. For in cases of an H-prediction based on an *explanans* containing an *antecedent* referring to a *future* interaction, there is also an asymmetry of assertibility between H-prediction and H-explanation in regard to the *explanans*. And second, the existence of biological properties which are *emergent* in the sense that even if all the laws were strictly *deterministic*, the occurrence of these properties could not have been *predicted* on the basis of any and all laws which could possibly have been discovered by humans in advance of the first known occurrence of the respective properties in question. Thus, evolutionary theory makes us familiar with past biological changes which were induced by prior past interactions, the latter being post-assertible on the basis of present records. And these past interactions can serve to explain the evolutionary changes in question. But the logical relation between *explanans* and *explanandum* furnishing this explanation is completely *time-symmetric*. Hence this situation makes for asymmetry only in the following innocuous sense: since corresponding future interactions cannot be rationally pre-asserted—there being no advance records of them—there is no corresponding pre-assertibility of those future evolutionary changes that will be effected by future interactions.

⁴ M. Scriven: "Explanation and Prediction in Evolutionary Theory," *op. cit.*, p. 477.

In an endeavor to establish the existence of an asymmetry damaging to Hempel's thesis on the basis of the account of a case of non-survival given in evolutionary theory, Scriven writes:

there are . . . good grounds [of inherent unpredictability] for saying that even in principle explanation and prediction do *not* have the same form. Finally, it is not in general possible to list all the exceptions to a claim about, for example, the fatal effects of a lava flow, so we have to leave it in probability form; this has the result of eliminating the very degree of certainty from the prediction that the explanation has, when we find the fossils in the lava.⁵

But all that the lava case entitled Scriven to conclude is that the merely *probabilistic* connection between the occurrence of a lava flow and the extinction of certain organisms has the result of depriving pre-assertibility of the very degree of certainty possessed by post-assertions here. Scriven is not at all justified in supposing that *predictive inferability* in this case lacks even an iota of the certainty that can be ascribed to the corresponding post-explanatory inferability. For wherein does the greater degree of certainty of the post-explanation reside? I answer: only in the assertibility of the *explanandum*, not in the character of the logical relation between the *explanans* (the lava flow) and the *explanandum* (fatalities on the part of certain organisms). What then must be the verdict on Scriven's contention of an asymmetry in the *certainty* of prediction and post-explanation in this context? We see now that this contention is vitiated by a confusion between the following two radically distinct kinds of asymmetry: First, a difference in the degree of certainty (categoricity) of our knowledge of the truth of the *explanandum* and of the claim of environmental unfitness made by the *explanans*, and second, a difference in the "degree of entailment," as it were, linking the *explanandum* to the *explanans*.

Very similar difficulties beset Scriven's analysis of a case of biological survival which is accounted for on the basis of environmental fitness. He says:

It is fairly obvious that no characteristics can be identified as contributing to "fitness" in all environments. . . . we cannot

⁵ *Ibid.*, p. 480.

predict which organisms will survive except in so far as we can predict the environmental changes. But we are very poorly equipped to do this with much precision." . . . However, these difficulties of prediction do not mean that the idea of fitness as a factor in survival loses all of its *explanatory* power. . . . animals which happen to be able to swim are better fitted for surviving a sudden and unprecedented inundation of their arid habitat, and in some such cases it is just this factor which explains their survival. Naturally we could have said in advance that *if* a flood occurred, they would be likely to survive; let us call this a hypothetical probability prediction. But hypothetical predictions do not have any value for actual predictions except in so far as the conditions mentioned in the hypothesis are predictable. . . . hence there will be cases where we can *explain why* certain animals and plants survived even when we could not have *predicted that* they would.⁷

There would, of course, be complete agreement with Scriven, if he had been content to point out in this context, as he does, that there are cases in which we can "explain why" but not "predict that." But he combines this correct formulation with the incorrect supposition that cases of post-explaining survival on the basis of fitness constitute grounds for an indictment of Hempel's thesis of symmetry. Let me therefore state the points of agreement and disagreement in regard to this case as follows. Once we recognize the ubiquitous role of *interactions* we can formulate the valid upshot of Scriven's observations by saying: insofar as future fitness and survival depend on future interactions which cannot be predicted from given information, whereas past fitness and survival depended on past interactions which *can* be retrodicted from that same information, there is

⁶ The environmental changes which Scriven goes on to cite are all of the nature of *interactions* of a potentially open system. And it is this common property of theirs which makes for their role in precluding the predictability of survival.

⁷ *Ibid.*, p. 478. In a recent paper "Cause and Effect in Biology" (*Science*, Vol. CXXXIV [1961], p. 1504), the zoologist E. Mayr overlooks the fallacy in Scriven's statement which we are about to point out and credits Scriven with having "emphasized quite correctly that one of the most important contributions to philosophy made by the evolutionary theory is that it has demonstrated the independence of explanation from prediction." And Mayr rests this conclusion among other things on the contention that "The theory of natural selection can describe and explain phenomena with considerable precision, but it cannot make reliable predictions."

an epistemological asymmetry between H-explanation and H-prediction in regard to the following: the assertibility both of the antecedent fitness affirmed in the *explanans* and of the *explanandum* claiming survival.

This having been granted as both true and illuminating, we must go on to say at once that the following considerations—which Scriven can grant only on pain of inconsistency with his account of asymmetry in the lava case—are no less true: the scientific inferability from a cause and hence our understanding of the *why* of survival furnished by an *explanans* which *does* contain the antecedent condition that the given animals are able to swim during a sudden, unprecedented inundation of their arid habitat is *not* one iota more probabilistic (i.e., *less* conclusive) in the case of a *future* inundation and survival than in the case of a *past* one. For if the logical nerve of intelligibility linking the *explanans* (fitness under specified kinds of inundational conditions) with the *explanandum* (survival) is only probabilistic in the future case, how could it possibly be any less probabilistic in the past case? It is evident that post-explanatory inductive inferability is entirely on a par here with predictive inferability from fitness as a cause. Why then does Scriven feel entitled to speak of “probability prediction” of *future* survival *without also* speaking of “probability explanation” of past survival? It would seem that his reason is none other than the *pseudo-contrast* between the *lack* of pre-assertibility of the *explanandum* (which is conveyed by the term “probability” in “probability prediction”) with the obtaining of post-explanatory inductive inferability of the *explanandum*. And this pseudo-contrast derives its plausibility from the tacit appeal to the bona fide asymmetry between the pre-assertibility and post-assertibility of the *explanandum*, an asymmetry which cannot score against the Popper-Hempel thesis.

II. *The Paresis Case.*

In a further endeavor to justify his repudiation of Hempel’s thesis, Scriven says:

we can explain but not predict whenever we have a proposition of the form “The only cause of X is A” (I)—for example, “The only cause of paresis is syphilis.” Notice that this is perfectly compatible with the statement that A is often not followed by

X—in fact, very few syphilitics develop paresis (II). Hence, when A is observed, we can predict that X is *more* likely to occur than without A, but still extremely unlikely. So, we must, on the evidence, still predict that it will *not* occur. But if it does, we can appeal to (I) to provide and guarantee our explanation. . . . Hence an event which cannot be predicted from a certain set of well-confirmed propositions can, if it occurs, be explained by appeal to them.⁸

In short, Scriven's argument is that although a past case of paresis can be explained by noting that syphilis was its cause, one cannot predict the future occurrence of paresis from syphilis as the cause. And he adds to this the following further comment:

Suppose for the moment we include the justification of an explanation or a prediction in the explanation or prediction, as Hempel does. From a general law and antecedent conditions we are then entitled to deduce that a certain event will occur in the future. This is the deduction of a prediction. From one of the propositions of the form the only possible cause of y is x and a statement that y has occurred we are able to deduce, not only that x must have occurred, but also the proposition the cause of y in this instance was x. I take this to be a perfectly sound example of deducing and explanation. Notice however, that what we have deduced is not at all a description of the event to be explained, that is we have not got an *explanandum* of the kind that Hempel and Oppenheim envisage. On the contrary, we have a specific causal claim. This is a neat way of making clear one of the differences between an explanation and a prediction; by showing the different kinds of proposition that they often are. When explaining Y, we do not have to be able to deduce that Y occurs, for we typically know this already. What we have to be able to deduce (if deduction is in any way appropriate) is that Y occurred *as a result of* a certain X, and of course this needs a very different kind of general law from the sort of general law that is required for prediction.⁹

⁸ M. Scriven: "Explanation and Prediction in Evolutionary Theory," *op. cit.*, p. 480.

⁹ M. Scriven: "Comments on Professor Grünbaum's Remarks at The Wesleyan Meeting," *Philosophy of Science*, Vol. XXIX (1962), pp. 173–74. For Scriven's most recent criticism of my views, see his "The Temporal Asymmetry of Explanations and Predictions," in: B. Baumrin (ed.) *Philosophy of Science* (New York: John Wiley and Sons; 1963), Vol. I, pp. 97–105.

I shall now show that Scriven's treatment of such cases as post-explaining paresis on the basis of syphilis suffers from the same defect as his analysis of the evolutionary cases: *Insofar as there is an asymmetry, Scriven has failed to discern its precise locus, and having thus failed, he is led to suppose erroneously that Hempel's thesis is invalidated by such asymmetry as does obtain.*

Given a particular case of paresis as well as the proposition that the only cause of paresis is syphilis—where a "cause" is understood here with Scriven as a "contingently necessary condition"—what can be inferred? Scriven maintains correctly that what follows is that both the paretic concerned had syphilis and that in his particular case, syphilis was the cause in the specified sense of "cause." And then Scriven goes on to maintain that his case against Hempel is established by the fact that we are able to assert that syphilis *did cause* paresis while *not* also being entitled to say that syphilis *will cause* paresis. But Scriven seems to have completely overlooked that our not being able to make both of these assertions does not at all suffice to discredit Hempel's thesis, which concerns the time-symmetry of the *inferability* of the *explanandum* from the *explanans*. The inadequacy of Scriven's argument becomes evident the moment one becomes aware of the reason for not being entitled to say that syphilis "will cause" paresis though being warranted in saying that it "did cause" paresis.

The sentences containing "did cause" and "will cause" respectively each make two affirmations as follows: First, the assertion of the *explanandum* (paresis) *per se*, and second, the affirmation of the obtaining of a causal relation (in the sense of being a contingently necessary condition) between the *explanans* (syphilis) and the *explanandum* (paresis). Thus, for our purposes, the statement "Syphilis *will cause* person Z to have paresis" should be made in the form "Person Z will have paresis *and* it will have been caused by syphilis," and the statement "Syphilis *did cause* person K to have paresis" becomes "Person K has (or had) paresis *and* it was caused by syphilis." And the decisive point is that in so far as a *past* occurrence of paresis can be inductively *inferred from* prior syphilis, so also a future occurrence of paresis can be. For the causal relation or connection between syphilis and paresis is incontestably time-symmetric:

precisely in the way and to the extent that syphilis *was* a necessary condition for paresis, it also *will be*! Hence the only bona fide asymmetry here is the record-based but innocuous one in the *assertibility* of the *explanandum* per se, but there is no asymmetry of *inferability* of paresis from syphilis. The former innocuous asymmetry is the one that interdicts our making the predictive assertion "will cause" while allowing us to make the corresponding post-explanatory assertion "did cause." And it is this fact which destroys the basis of Scriven's indictment of Hempel's thesis. For Hempel and Oppenheim did not maintain that an *explanandum* which can be post-asserted can always also be pre-asserted; what they did maintain was only that the *explanans* never post-explains any better or more conclusively than it implies predictively, there being complete symmetry between post-explanatory *inferability* and predictive *inferability* from a given *explanans*. They and Popper were therefore fully justified in testing the adequacy of a proffered *explanans* in the social sciences on the basis of whether the post-explanatory inferability of the *explanandum* which was claimed for it was matched by a corresponding predictive inferability, either inductive or deductive as the case may be.

What is the force of the following comment by Scriven: in the post-explanation of paresis we do not need to infer the *explanandum* from the *explanans* à la Hempel and Oppenheim, because we know this already from prior records (observations) of one kind or another; what we do need to infer instead is that the *explanandum*-event occurred *as a result of* the cause (necessary condition) given by the *explanans*, an inference which does not allow us to *predict* (i.e., pre-assert) the *explanandum*-event? This comment of Scriven's proves only that here there is record-based post-assertibility of paresis but no corresponding pre-assertibility.

In short, Scriven's invocation of the paresis case, just like his citation of the cases from evolutionary theory, founders on the fact that he has confused an *epistemological* asymmetry with a *logical* one. To this charge, Scriven has replied irrelevantly that he has been at great pains in his writings—as for example in his discussion of the barometer case which I shall discuss below—to distinguish valid arguments based on true premises which do

qualify as scientific explanations from those which do not so qualify. This reply is irrelevant, since Scriven's caveat against identifying (confusing) arguments based on true premises which are both valid and explanatory with those which are valid without being explanatory does not at all show that he made the following crucial distinction here at issue: the distinction between (1) a difference (asymmetry) in the assertibility of either a conclusion (*explanandum*) or a premise (*explanans*), and (2) a difference (asymmetry) in the inferability of the *explanandum* from its *explanans*. Although the distinction which Scriven does make cannot serve to mitigate the confusion with which I have charged him, his distinction merits examination in its own right.

To deal with it, I shall first consider examples given by him which involve *non-predictive* valid deductive arguments to which he denies the status of being explanatory arguments. And I shall then conclude my refutation of Scriven's critique of Hempel's thesis by discussing the following paradigm case of his: the deductively valid *predictive* inference of a storm from a sudden barometric drop, which he adduces in an endeavor to show that such a valid deductive inference could not possibly qualify as a *post-explanation* of a storm.

It would be agreed on all sides, I take it, that no *scientific understanding* is afforded by the deduction of an *explanandum* from itself even though such a deduction is a species of valid inference. Hence it can surely be granted that the class of valid deductive arguments whose conclusion is an *explanandum* referring to some event or other is wider than the class of valid deductive arguments affording scientific understanding of the *explanandum*-event. But it is a quite different matter to claim, as Scriven does, that no *scientific understanding* is provided by those valid deductive arguments which ordinary usage would not allow us to call "explanations." For example, Scriven cites the following case suggested by S. Bromberger and discussed by Hempel:¹ the height of a flagpole is deducible from the length of its shadow and a measurement of the angle of the sun taken in conjunction with the principles of geometrical optics, but the height of the flagpole could not thereby be said to have been

¹ Cf. C. G. Hempel: "Deductive-Nomological vs. Statistical Explanation," *op. cit.*, Sec. 4.

"explained." Or take the case of a rectilinear triangle in physical space for which Euclidean geometry is presumed to hold, and let it be given that two of the angles are 37° and 59° respectively. Then it can be deductively inferred that the third angle is one of 84° , but according to Scriven, this would not constitute an explanation of the magnitude of the third angle.

Exactly what is shown by the flagpole and angle cases concerning the relation between valid deductive arguments which furnish scientific understanding and those which, according to ordinary usage, would qualify as "explanations"? I maintain that while differing in one respect from what are usually called "explanations," the aforementioned valid deductive arguments yielding the height of the flagpole and the magnitude of the third angle provide scientific understanding no less than "explanations" do. And my reasons for this contention are the following.

In the flagpole case, for example, the *explanandum* (stating the height of the flagpole) can be deduced from two different kinds of premises: First, an *explanans* of the type familiar from geometrical optics and involving laws of coexistence rather than laws of succession, antecedent events playing no role in the *explanans*, and second, an *explanans* involving causally antecedent events and laws of succession and referring to the temporal genesis of the flagpole as an artifact. But is this difference between the kinds of premises from which the *explanandum* is deducible a basis for claiming that the coexistence-law type of *explanans* provides less *scientific understanding* than does the law-of-succession type of *explanans*? I reply: certainly not. And I hasten to point out that the difference between *pre*-axiomatized and axiomatized geometry conveys the measure of the scientific understanding provided by the geometrical account given in the flagpole and angle cases on the basis of laws of coexistence. But is it not true after all that ordinary usage countenances the use of the term "explanation" only in cases employing causal antecedents and laws of succession in the *explanans*? To this I say: this *terminological* fact is as unavailing here as it is philosophically unedifying.

Finally, we turn to Scriven's citation of cases of deductively valid predictive inferences which, in his view, invalidate Hempel's thesis because they could not possibly also qualify as post-explanations.

III. *The Barometer Case.*

Scriven writes:

What we are trying to provide when making a prediction is simply a claim that, at a certain time, an event or state of affairs will occur. In explanation we are looking for a cause, an event that not only occurred earlier but stands in a special relation to the other event. Roughly speaking, the prediction requires only a correlation, the explanation more. This difference has as one consequence the possibility of making predictions from indicators other than causes—for example, predicting a storm from a sudden drop in the barometric pressure. Clearly we could not say that the drop in pressure in our house caused the storm: it merely presaged it. So we can sometimes predict what we cannot explain.²

Other cases of the barometer type are cases such as the presaging of mumps by its symptoms and the presaging of a weather change by rheumatic pains.

When we make a predictive inference of a storm from a sudden barometric drop, we are inferring an effect of a particular cause from another (earlier) presumed effect of that same cause. Hence the inference to the storm is not from a cause of the storm but only from an indicator of it. And the law connecting sudden barometric drops to storms is therefore a law affirming only an indicator type of connection rather than a causal connection.

The crux of the issue here is whether we have no scientific understanding of phenomena on the strength of their deductive inferability from indicator laws (in conjunction with a suitable antecedent condition), scientific understanding allegedly being provided only by an *explanans* making reference to one or more causes. If that were so, then Scriven could claim that although the mere inferability of particular storms from specific sudden barometric drops is admittedly time-symmetric, there is no time-symmetry in positive scientific understanding. It is clear from the discussion of the flagpole case that the terminological practice of restricting the term "explanation" though not the term "prediction" to cases in which the *explanans* makes reference to a partial or total cause rather than to a mere indicator cannot

² M. Scriven: "Explanation and Prediction in Evolutionary Theory," *op. cit.*, p. 480.

settle the questions at issue, which are: would the type of argument which yields a *prediction* of a future *explanandum*-event (storm) from an indicator-type of premise furnish any scientific understanding, and, if so, does this type of argument provide the same positive amount of scientific understanding of a corresponding *past* event (storm)?

These questions are, of course, not answered in the negative by pointing out correctly that the law connecting the cause of the storm with the storm can serve as a *reason* for the weaker indicator-law. For this fact shows only that the causal law can account for both the storm and the indicator law, but it does not show that the indicator law cannot provide any scientific understanding of the occurrence of particular storms. To get at the heart of the matter, we must ask what distinguishes a causal law from an indicator law such that one might be led to claim, as Scriven does, that subsumption under indicator laws provides no scientific understanding at all, whereas subsumption under causal laws does.

Let it be noted that a causal law which is used in an *explanans* and is not itself derived from some wider causal law is fully as *logically contingent* as a mere indicator law which is likewise not derived from a causal law but is used as a premise for the deduction of an *explanandum* (either predictively or postdictively, i.e., H-explanatorily). Why then prefer (predictive or postdictive) subsumption of an *explanandum* under a causal law to subsumption under a mere indicator law? The justification for this preference would seem to lie not merely in the greater generality of the causal law; it also appears to rest on the much larger variety of empirical contingencies which must be ruled out in the *ceteris paribus* clause specifying the relevant conditions under which the indicator law holds, as compared to the variety of such contingencies pertaining to the corresponding causal law.³ But this difference in both generality and in the variety of contingencies does not show that the indicator-law

³ For example, in addition to all the things that can interfere with the materializing of a storm when its avowed total cause (a pressure drop over a wide area) is presumed to be present, there are a variety of further contingencies such that the occurrence of any one of them will eventuate in the *absence* of a storm when we observe a sudden barometric drop at one place: this sudden drop may arise from a *narrowly local* pressure-lowering cause (device) in the immediate spatial vicinity of the barometer. And

provides no scientific understanding of particular phenomena subsumable under it; it shows only, so far as I can see, that one might significantly speak of degrees of scientific understanding. And this conclusion is entirely compatible with the contention required by the symmetry thesis that the barometric indicator law furnishes the same positive amount of scientific understanding of a past storm as of a future one predicted by it.

I believe to have shown, therefore, that with respect to the symmetry thesis, *Hempel ab omni naevo vindicatus*.⁴

(D) THE CONTROVERSY BETWEEN MECHANISM AND TELEOLOGY.

The results of our discussion of the temporal asymmetry of recordability have a decisive bearing on the controversy between mechanism and teleology.

hence, in that case, the sudden barometric drop would not at all betoken the presence of a pressure drop over an area sufficiently large to eventuate in a storm. Similarly, the occurrence of what appear to be symptoms of mumps may not betoken the presence of the filter-passing virus which can cause this disease; instead these presumed symptoms may arise from any one of a number of other kinds of causes, no one of which would issue in mumps. These other kinds of causes must be ruled out, *in addition* to those entering the *ceteris paribus* clause of the law relating the mumps to *its cause*, if the *symptoms* of mumps are to serve as a reliable basis for inferring the subsequent onset of this disease.

Although the distinction between a causal law (C-law) and an indicator law (I-law) seems to be sufficiently clear in the examples adduced by Scriven, Professor Richard Rudner has suggested in private correspondence that, in view of the well-known difficulties besetting the characterization of a C-law as such, we should be leery of supposing that a generally clear and tenable distinction between C-laws and I-laws has been drawn. In the barometer and mumps cases, we distinguish the initial total cause from a mere indicator by pointing out that the indicator is itself a partial effect of the initial total cause. If the difficulties besetting the general characterization of a C-law are actually such that this distinguishing criterion will not do and such that no other viable criterion seems to be in sight, then my remarks about our preference for subsumption of an *explanandum* under a C-law rather than an I-law would lose generality and would have to be restricted to specific examples like those cited by Scriven. It is clear that Scriven's argument here depends for its very statement on the tenability of the distinction between C-laws and I-laws. And if that distinction were indeed fundamentally untenable, this fact alone would suffice to refute Scriven's argument.

⁴ Believing (incorrectly) to have cleansed Euclid of all blemish, G. Saccheri (1667–1733) published a book in 1733 under the title *Euclides ab omni naevo vindicatus*.

By mechanism we understand the philosophical thesis that all explanation must be *only a tergo*, i.e., that occurrences at a time t can be explained *only* by reference to *earlier* occurrences and *not also* by reference to later ones.⁵ And by teleology we understand a thesis which is the contrary rather than the contradictory of mechanism: all phenomena belonging to a specified domain and occurring at a time t are to be understood by reference to *later* occurrences only. We note that, thus understood, mechanism and teleology can both be false.

During our post-Newtonian epoch there is a misleading incongruity in using the term "mechanism" for the thesis of the monopoly of *a tergo* explanations. For in the context of the *time-symmetric* laws of Newton's mechanics, the given state of a closed mechanical system at a time t can be inferred from a state *later* than t (i.e., *retrodicted*) no less than the given state can be inferred from a state *earlier* than t (i.e., *predicted*). Instead of furnishing the prototype for mechanistic explanation in the philosophical sense, the phenomena described by the time-symmetric laws of Newton's mechanics constitute a domain with respect to which both mechanism and teleology are false, thereby making the controversy between them a pseudo-issue. More generally, that controversy is a pseudo-issue with respect to any domain of phenomena constituted by the evolution of closed systems obeying time-symmetric laws, be they deterministic or statistical.

But there is indeed a wide class of phenomena with respect to which mechanism is true. And one may presume that tacit reference to this particular class of phenomena has conferred plausibility on the thesis of the *unrestricted* validity of mechanism: traces or marks of interaction existing in a system which is essentially closed at a time t are accounted for scientifically by *earlier interactions* or *perturbations* of that system—which are called "causes"—and not by later interactions of the system. Thus, we

⁵ A weaker version of mechanism might be the thesis that all phenomena can be understood by reference to earlier occurrences, a thesis which *allows* that understanding might be provided by reference to later occurrences as well. This weaker and less influential version of mechanism, though likewise incompatible with teleology, is not, however, the one whose assessment is illuminated by the asymmetry of recordability. Hence I discuss only the stronger, though more vulnerable, version of mechanism here.

explain a scar on a person by noting that he did sustain an injury in the past, not by claiming that he will suffer injury in the future.

In view of the demonstrated *restricted* validity of mechanism, we must therefore deem the following statement by H. Reichenbach as too strong: "We conclude: If we define the direction of time in the usual sense, there is no finality, and only causality is accepted as constituting explanation."⁶

⁶ H. Reichenbach: *DT, op. cit.*, p. 154.

Chapter 10

See Append. §31

IS THERE A "FLOW" OF TIME OR TEMPORAL "BECOMING"?

It is clear that the anisotropy of time resulting from the existence of irreversible processes consists in the mere structural differences between the two opposite senses of time but provides no basis at all for singling out *one* of the two opposite senses as "*the* direction" of time. Hence the assertion that irreversible processes render time anisotropic is *not at all* equivalent to such statements as "time flows one way."

Thus we must clearly distinguish the *anisotropy* of physical time from the feature of common sense (psychological) time which is rendered by such terms as "the transiency of the Now" or "becoming" and by such metaphors as the "flux," "flow" or "passage" of time. And our concern in the present chapter is to determine the factual credentials, if any, of the concepts which are represented by these terms.

Since the instants of anisotropic time are ordered by the relation "earlier than" no less than by the converse relation "later than," the anisotropy of time provides no warrant at all for singling out the "later than" sense as "the" direction of time. Instead, the inspiration for speaking about "the" direction of time derives from the supposition that there is a transient "now" or "present" which can be claimed to shift so as to single out the future direction of time as the sense of its "advance."

The transient or shifting division of the time continuum into the *past* and the *future* depends on the transient Now and is *not*

furnished solely by the "static" relation of "earlier than" or its converse "later than" with respect to which we formulated the anisotropy of time. The obtaining of the relation "earlier than" or "later than" between two physical events or states does not, of course, depend at all on the transient Now.¹ Thus the year 1910 is earlier than the year 1920, no less than 1950 is earlier than 1970, and than 1970, in turn, is earlier than 2850. And yet, at this writing in 1962, which is at the focus of my immediate experience, the events of 1970 and 2850 belong to the future while those of the other years we mentioned belong to the past such that the Now within 1962 on which this classification depends is not an arbitrarily chosen time of reference. For the past is the class of events earlier than those events which constitute the Now in the sense that the past is constituted by the events which "no longer exist" while those of the Now or present "actually exist." And the future is correspondingly the set of events that are later than now in the sense that they are yet to "acquire existence," as it were, or to "come *into* being." Thus, writing about himself in 1925, Reichenbach speaks of having "the feeling that *my* existence is a reality, while *Plato's* life merely still casts its shadows onto reality" and declares that we are unable to escape "the compulsion which distinguishes for us a Now-point in an absolute way as the experience of the divide between past and future."²

The transience of the Now is a feature of psychological (and common sense) time in the sense that there is a *diversity* of the Now-contents of immediate awareness. Hence it is a matter of fact that the Now "shifts" in conscious awareness to the extent that there is a *diversity* of the Now-contents, and it is likewise a fact that the Now-contents are temporally *ordered*. But since these diverse Now-contents are ordered with respect to the relation "earlier than" no less than with respect to its converse "later

¹ This independence *cannot*, of course, be gainsaid by the following consideration from the special theory of relativity (cf. Chapter Twelve): in the case of those and only those particular pairs of events which are *not* linkable by causal chains, the usual choice of a definition of simultaneity made in that theory leads to a dependence *on the inertial system* of the conventional ascription of the relations "later than" and "simultaneous with."

² H. Reichenbach: "Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft," *op. cit.*, pp. 133-75, esp. p. 140.

than," it is a mere *tautology* to say that the Now shifts *from* earlier *to* later. For this metaphorical affirmation of shifting in *the future direction along the time-axis* tells us no more than that later Nows are later than earlier ones, just as earlier Nows are earlier than later ones! By the same token, the assertion that the "flow" of time is *unidirectional* is a tautology, as is the claim that time "flows" from the past to the future. The factual emptiness of the latter formulations must not tempt us into overlooking the following: a *non-directional* or directionally *neutral* claim of the transiency of the Now and of the temporal order of the various "Now"-contents does codify factual truths pertaining to psychological (common sense) time. And we shall see in some detail later on that because of its inherent dependence on consciousness, the transiency of the Now is not also a feature of physical time.

These considerations now enable us to evaluate the following contentions by Reichenbach:

We have to distinguish here between two problems. First, the procedure described leads to an *order* of time, in the same sense in which the points on a line are ordered. Such a series of points has two directions, neither of which has any distinguishing characteristic. Temporal order, too, has two directions, the direction to earlier and the direction to later events, but in this case one of the two directions has a distinguishing characteristic: time flows from the earlier to the later event. Time therefore represents not only an ordered series generated by an asymmetrical relation, but is also *unidirectional*. This fact is usually ignored. We often say simply: the direction from earlier to later events, from cause to effect, is the direction of the progress of time. However, in this form the assertion is empty unless we specify what "progress of time" means. In the same fashion we could say that the points on a line progress from left to right; but this assertion is empty, since the progress of points means here nothing but the progressing in the selected direction. When we speak about the progress of time, in contrast, we intend to make a synthetic assertion which refers both to an immediate experience and to physical reality. This particular problem can only be solved if we can formulate the content of the assertion more precisely. We shall leave this problem for the time being and content ourselves with the con-

clusion that the direction which we have defined as earlier-later is the same direction as that of the progress of time. For the problems dealt with in the theory of relativity it suffices that there exists a serial *order* of time, i.e., that we can distinguish between two directions which are opposite to each other.³

We saw that the anisotropy of time lies in the fact that *each* of the two opposite directions of time has a distinguishing characteristic. But Reichenbach believes that the transiency of the Now entitles him to maintain the following: "one of the two directions has a distinguishing characteristic: time flows from the earlier to the later event" and is "*unidirectional*" in this sense. Discerning clearly that the charge of tautology might be leveled against the assertion that the direction from earlier to later is "the direction of the progress of time," Reichenbach contends that this assertion can be synthetic. And his reason is that the assertion renders the content of "an immediate experience" as well as an objective feature of "physical reality." Deferring until later our appraisal of the relevance of the transient Now to *physical* as distinct from common sense (psychological) time, we see that Reichenbach overlooked a fundamental point here: it is only the obtaining of a *diversity* of Now-contents which is a matter of fact, but not the allegedly "*unidirectional*" character of "the progress of time." For, as we saw earlier, the factuality of the diversity of the Now-contents does not suffice to give synthetic content to the assertion that time progresses unidirectionally from earlier to later. And Reichenbach's allegation here of the synthetic character of this assertion is avowedly left unsubstantiated by him here.

We shall now turn to a critical analysis of those of his writings in which he attempted to justify his cited contention that a Now and its associated transient division between the past and the future is a feature of *physical* time or "physical reality" no less than of psychological (common sense) time.

Having noted that the concept of the purported unidirectional progress of time has not found a place within the theory of relativity because "it suffices that there exists a serial *order* of time"⁴

³ H. Reichenbach: *PST*, *op. cit.*, pp. 138-39. Reprinted through permission of the publisher.

⁴ *Ibid.*, p. 139.

for the problems dealt with in that theory, Reichenbach deems that theory's Minkowskian world picture incomplete as follows:

A topology of time can be constructed in which the basic concepts "earlier," "later" and "simultaneous" are defined. But what could *not* be solved in this way so far is the problem of the "now."

What does "now" mean? Plato lived before me, and Napoleon VII will live after me. But which one of these three lives *now*? I undoubtedly have a clear feeling that *I* live now. But does this assertion have an objective significance beyond my subjective experience?⁵

. . . In the condition of the world, a cross-section called the present is distinguished; the "now" has an objective significance. Even when no human being is alive any longer, there is a "now"; the "present state of the planetary system" is then just as much a determinate specification as "the state of the planetary system at the time of the birth of Christ."

In the four-dimensional picture of the world, such as used by the theory of relativity, there is no such distinguished cross-section. But this is due only to the fact that an essential content is omitted from this picture.⁶

That the theory of relativity does not make any allowance for the *transient* Now of common sense time is indeed correct. For the "Now" in the "Here-Now" of the Minkowski diagram designates no more than a kind of *arbitrary* zero or origin of temporal coordinates: we can make use of the Minkowski diagram at noon on June 1, 1962 to let "Here-Now" designate a certain event occurring after the extinction of the sun. And the "Absolute Past" and the "Absolute Future" are no more than the set of events respectively absolutely earlier and later than the event *arbitrarily* designated as the "Here-Now." Instead of allowing for the transient division of time into the past and future by the shifting Now of experienced time, the theory of relativity conceives of events as simply being and sustaining relations of earlier and later, but not as "coming *into* being": we conscious organisms then "come across" them by "entering" into their absolute future,

⁵ H. Reichenbach: "Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft," *op. cit.*, p. 139.

⁶ *Ibid.*, p. 141.

as it were. And upon experiencing their immediate effects, we regard them as "taking place" or "coming into being."⁷

Like Eddington⁸ before him and G. J. Whitrow⁹ after him, Reichenbach supposed that the exclusion of the concept of the unidirectional progress of time by a physical theory is attributable to the *deterministic* character of that theory. And believing—quite erroneously as we shall see—that an *indeterministic* physics can provide a *physical* basis for the transient Now, he attempted to find an objective physical basis for the present or Now in his paper of 1925¹ as follows: he gave a *probabilistic* interpretation of causality according to which the past is "objectively determined," while the future is "objectively undetermined" in virtue of the non-existence of a complete set of partial causes, a knowledge of which would render our predictions certain. In this context he conceived the present *without* the inadmissible use of absolute simultaneity as the class of events not causally connectible with the particular "now." But (as is evident from the world-line diagram in Chapter 12, p. 354), this characterization yields a conception of the present which differs from the "Now" of conscious experience in the following essential respect: a given event E_2 at a point P_2 in space will remain simultaneous for an observer at the distant point P_1 *throughout* a time interval constituted by a continuum of events having a space-like separation from E_2 ! It is because Reichenbach maintained that on the basis of determinism, "the morrow has already occurred today in the same sense as yesterday has," thus purportedly making nonsense of all our planning, that he rejected determinism in 1925 just before the full advent of quantum mechanics and sought a physical basis for the present and thereby for becoming. Just before his death in 1953, he argued²

⁷ Cf. A. S. Eddington: *The Nature of the Physical World*, *op. cit.*, p. 68, and *Space, Time and Gravitation*, *op. cit.*, p. 51. See also E. Cassirer: *Zur Einsteinschen Relativitätstheorie* (Berlin: Bruno Cassirer Verlag; 1921), pp. 120–21.

⁸ A. S. Eddington: *Space, Time and Gravitation*, *op. cit.*, p. 51.

⁹ G. J. Whitrow: *The Natural Philosophy of Time*, *op. cit.*, p. 295.

¹ H. Reichenbach: "Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft," *op. cit.*, pp. 141–43.

² H. Reichenbach: *DT*, *op. cit.*, pp. 211–24 and "Les Fondements Logiques de la Mécanique des Quanta," *op. cit.*, pp. 154–57.

that the micro-indeterminism of quantum mechanics is not an ephemeral *pis-aller* of present-day physical theory and extended his early views by attempting to utilize the indeterminacies of quantum mechanics. He writes:

Let us suppose that consecutive measurements are made alternately of two noncommuting [i.e., complementary] quantities. One will obtain a series of macroscopic events which one cannot predict, but which one can record. This series provides us with a clear distinction between the past and the future: the past is determined, but the future is not. . . . The analysis of classical physics has shown us that one can record the past but not the future. The combination of this result with Heisenberg's uncertainty leads us to the consequence that one can know the past but that one cannot predict the future. . . .

. . . Modern science . . . furnishes us with precisely the difference between the past and the future, which Laplace's physics could not recognize.

To be sure, Boltzmann's physics, if coupled with the hypothesis of the branch structure, yields a certain structural difference between the past and the future. . . . But while this difference enabled us to distinguish between the past and the future, it was not associated with a difference in determination: although one cannot record the future, one could predict it on the basis of the totality of causes. Thus, one cannot call the future undetermined. . . .

It is no longer that way in quantum physics. . . . Here is the difference: there are future facts which cannot possibly be predicted, whereas there are no past facts which it would be impossible to know. In principle, they can always be recorded. . . .

The distinction between the indeterminateness ("l'indéterminisme") of the future and the determinateness ("détermination") of the past has, in the final analysis, found expression in the laws of physics. . . . The concept of "becoming" acquires significance in physics: the present, which separates the future from the past, is the moment at which that which was undetermined becomes determined, and "becoming" has the same meaning as "becoming determined." . . .

. . . The term "determination" denotes a relation between two situations A and B; the situation A does or does not determine the situation B. It is meaningless to say that the situation B, considered by itself, is determined. If we say that the past is determined or that the future is undetermined, it is tacitly

understood that we are relating this to the present situation; it is with respect to "now" that the past is determined and that the future is not.³

Reichenbach's view here is shared by Eddington, who writes:

The division into past and future (a feature of time-order which has no analogy in space-order) is closely associated with our ideas of causation and free-will. In a perfectly determinate scheme the past and future may be regarded as lying mapped out—as much available to present exploration as the distant parts of space. Events do not happen; they are just there, and we come across them. "The formality of taking place" is merely the indication that the observer has on his voyage of exploration passed into the absolute future of the event in question; and it has no important significance.⁴

In the same vein, the astronomer H. Bondi contends that "In a theory with indeterminacy . . . the passage of time transforms statistical expectation into real events."⁵ And G. J. Whitrow claims⁶ that "There is indeed a profound connection between the reality of time and the existence of an incalculable element in the universe." If Reichenbach, Eddington, Bondi, Whitrow and others had merely maintained that indeterminacy makes for our human inability to know in advance of their actual occurrence what particular kinds of events will in fact materialize, then, of course, there could be no objection. But I take them to have claimed that the *existential* status of future events in an indeterministic world is that of *coming into being* with time, whereas in a deterministic world it is one of simply being.

I believe that the issue of determinism *vs.* indeterminism is *totally irrelevant* to whether becoming is a significant attribute of the time of physical nature independently of human consciousness. And I wish to explain now why I regard the thesis of Reichenbach, Eddington, Bondi, Whitrow, and of many others that indeterminism confers flux onto physical time as untenable. I have given my reasons for likewise rejecting Reichenbach's

³ H. Reichenbach: "Les Fondements Logiques de la Mécanique des Quanta," *op. cit.*, pp. 154–57.

⁴ A. S. Eddington: *Space, Time and Gravitation*, *op. cit.*, p. 51.

⁵ H. Bondi: "Relativity and Indeterminacy," *Nature*, Vol. CLXIX (1952), p. 660.

⁶ G. J. Whitrow: *The Natural Philosophy of Time*, *op. cit.*, p. 295.

further claim that "The paradox of determinism and planned action is a genuine one"⁷ in other publications.⁸

In the indeterministic quantum world, the relations between the sets of measurable values of the state variables characterizing a physical system at different times are, in principle, not the one-to-one relations linking the states of classically behaving closed systems. But this holds for a given state of a physical system and its absolute future quite independently of whether that state occurs at midnight on December 31, 1800 or at noon on March 1, 1984. Moreover, if we consider *any one* of the temporally successive regions of space-time, we can assert the following: the events belonging to its particular absolute past could be (more or less) uniquely specified in records which are a part of that region, whereas its particular absolute future is thence quantum mechanically unpredictable. Accordingly, every "now," be it the "now" of Plato's birth or that of Reichenbach's, always constitutes a divide in Reichenbach's sense between its own recordable past and its unpredictable future, thereby satisfying Reichenbach's definition of the "present." But this fact is fatal to his avowed aim of providing a physical basis for a "unique," transient "now" and thus for "becoming."⁹ Reichenbach's recent characterization of the determinacy of the past as recordability as opposed to the quantum mechanical indeterminacy of the future can therefore not serve to vindicate his conception of becoming any more than did his paper of 1925,¹ which was penetratingly criticized by Hugo Bergmann as follows:

Thus, according to Reichenbach, a cross-section in the state of the world is distinguished from all others; the now has an

See Append. § 31

⁷ H. Reichenbach: *DT*, *op. cit.*, p. 12.

⁸ Cf. A. Grünbaum: "Causality and the Science of Human Behavior," *American Scientist*, Vol. XL (1952), pp. 665-76 [reprinted in: H. Feigl and M. Brodbeck (eds.) *Readings in the Philosophy of Science* (New York: Appleton-Century-Crofts, Inc.; 1953), pp. 766-78]; "Das Zeitproblem," *Archiv für Philosophie*, Vol. VII (1957), pp. 203-206; "Complementarity in Quantum Physics and its Philosophical Generalization," *The Journal of Philosophy*, Vol. LIV (1957), pp. 724-27, and "Science and Man," *Perspectives in Biology and Medicine*, Vol. V (1962), pp. 483-502. See also J. J. C. Smart's telling criticisms [*Philosophical Quarterly*, Vol. VIII (1958), esp. p. 76] of Reichenbach's contention that we can "change the future" but not the past.

⁹ This aim is stated by him in "Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft," *op. cit.*, pp. 139-42.

¹ *Ibid.*

objective significance. Even when no man is alive any longer, there is a now. "The present state of the planetary system" would even then be just as precise a descriptive phrase as "the state of the planetary system in the year 1000."

Concerning this definition one must ask: Which now is intended, if one says: the present state of the planetary system? That of the year 1800 or 2000 or which other one? Reichenbach's reply is: the now is the threshold of the transition from the state of indeterminacy to that of determinacy. But (if Reichenbach's indeterminism holds) this transition has *always* occurred and will always occur. And if the rejoinder would be: the indeterminacy of the year 1800 has already been transformed into a determinacy, then one must ask: For whom? Evidently for us, for the present, for our now. Accordingly, this definition by Reichenbach seems to refer after all to a now which it must first define. What is the objective difference between the now of the year 1800 and the now of the present instant? The answer must be: now is the instant of the transition from indeterminacy to determinacy, that is, one explains the present now . . . by reference to itself.

. . . Reichenbach writes: The problem can be formulated as the question concerning the difference between the past and the future. For determinism, there is no such difference. . . . But the reproach which Reichenbach directs at determinism here should be aimed not at it but at the world view of physics, which does not take cognizance of any psychological categories, for which there is no "I," . . . a concept which is inextricably intertwined with the concept "now." Even those who regard the supplanting of determinism by indeterminism as admissible, as we do, will not be willing to admit that the concept of "now" can be assigned a legitimate place within indeterministic physics. Even if one assumes—as we wish to do along with Reichenbach—that the future is not uniquely determined by a temporal cross-section, one can say only that this indeterminacy prevails just as much for Plato as for myself and that I cannot decide by physical means who is living "now." For the difference is a psychological one.

. . . "Now" is the temporal mode of the experiencing ego.²

² H. Bergmann: *Der Kampf um das Kausalgesetz in der jüngsten Physik*, *op. cit.*, pp. 27–28. Wilfrid Sellars has independently developed the basis for similar criticisms of the alleged connection between indeterminism and becoming as part of his penetrating study of a complex of related issues: cf. W. Sellars's "Time and The World Order," in *Minnesota Studies in the Philosophy of Science*, Vol. III, *op. cit.*, pp. 527–616.

I maintain with Bergmann that the transient now with respect to which the distinction between the past and the future of common sense and psychological time acquires meaning has no relevance at all to the time of physical events, because it has no significance at all apart from the egocentric perspectives of a *conscious* (human) organism and from the immediate experiences of that organism. If this contention is correct, then *both in an indeterministic and in a deterministic world, the coming into being or becoming of an event, as distinct from its merely being, is thus no more than the entry of its effect(s) into the immediate awareness of a sentient organism (man)*. For what is the difference between these two worlds in regard to the determinateness of future events? The difference concerns only the type of functional connection linking the attributes of the future events to those of present or past events. But this difference does *not* make for a precipitation of future events into existence in a way in which determinism does not. Nor does indeterminacy make for any difference whatever at any time in regard to the *attribute-specificity* of the future events themselves. For in either kind of universe, it is a fact of logic that what will be, will be. The result of a future quantum mechanical measurement may not be definite prior to its occurrence in relation to earlier states, and thus our prior knowledge of it correspondingly cannot be definite. But as an event, it is as fully attribute-definite and occurs just as a measurement made in a deterministic world does. The belief that in an indeterministic world, the future events come into being or become actual or real with the passage of time would appear to confuse two quite different things: (1) the *epistemological* precipitation of the actual event-properties of future events out of the wider matrix of the possible properties allowed by the quantum-mechanical probabilities, and (2) an existential coming into being or becoming actual or real. Only the *epistemological* precipitation is affected by the passage of time through the transformation of a statistical expectation into a definite piece of information. But this does *not* show that in an indeterministic world there is any kind of *precipitation into existence* or *coming into being* with the passage of time. And even in a deterministic world, the effects of physical events come into our awareness at a certain time and *in that sense* can be thought of as coming into being.

Bergmann's demonstration above that an indeterminist universe fails to define a *non-psychological* objective transient now can be extended in the following sense to justify his contention that the concept "now" involves features peculiar to consciousness: the "flux of time" or transiency of the "now" has a meaning only in the context of the egocentric perspectives of *sentient* organisms and does *not also* have relevance to the relations between purely inanimate individual recording instruments and the environmental physical events they register, as Reichenbach claims. For what can be said of every state of the universe can also be said, *mutatis mutandis*, of every state of a given inanimate recorder. Moreover, the irrelevance of the transient now to the accretion of time-tagged marks or *traces* on an *inanimate* recording tape also emerges from William James's and Hans Driesch's correct observation that a simple isomorphism between a succession of *brain traces* and a succession of states of awareness does not explain the now-contents of such psychological phenomena as melody awareness. For the hypothesis of isomorphism of traces and states of awareness renders only the succession of states of awareness but not the *instantaneous awareness of succession*,³ which is an essential ingredient of the meaning of "now": the now-content, when viewed as such in awareness, includes an awareness of the order of succession of events in which the occurrence of that awareness constitutes a *distinguished element*. And the transiency of the now or the "flux" of time arises from the *diversity* of the now-contents having the latter attributes: there are striking differences in the *membership* of the set of remembered (recorded) and/or forgotten events of which we have instantaneous awareness.

I cannot see, therefore, that Reichenbach is justified in considering the accretion of time-tagged marks or traces on an *inanimate* recording tape so as to form an expanding spatial series as illustrating the "flux" of time. Thus, Bergmann's exclusively psychologistic conception of this flux or becoming must be upheld against Reichenbach: the flux depends for its very *existence* on the perspectival role of consciousness, since the coming *into* being (or becoming) of an event is no more than

³ Cf. W. James: *The Principles of Psychology*, *op. cit.*, pp. 628–29, and H. Driesch: *Philosophische Gegenwartsfragen* (Leipzig: E. Reinicke; 1933), pp. 96–103.

the entry of its effect(s) into the immediate awareness of a sentient organism (man).

We saw earlier that the locution "Here-Now" of the relativistic Minkowski diagram does not commit that entirely non-psychological theory to the transient now encountered in common sense time. Hence the purely physical character of the special theory of relativity *cannot* be adduced to show that the transient now is relevant to physical time, i.e., it cannot be adduced to refute our claim of the dependence of the "now" and, correlatively, of the transient division of the time continuum into "past" and "future," on the perspectival role of consciousness.⁴

It was none other than the false assumption that "flux" must be a feature of physical no less than of psychological (common sense) time that inspired Henri Bergson's misconceived polemic against the mathematical treatment of motion, which he unfoundedly charged with having erroneously *spatialized* time by a description which leaves out the flux of becoming and renders only the "static" relations of earlier and later.⁵

Hermann Weyl has given a metaphorical rendition of the dependence of coming into being on consciousness by writing: "The objective world simply *is*, it does not *happen*. Only to the gaze of my consciousness, crawling⁷ upward along the life- [i.e., world-] line of my body, does a section of this world come to life as a fleeting image in space which continuously changes in time." This poetic but sound declaration has given rise to serious misunderstandings, as shown by the following objection from Max Black:

⁴ Neither does Minkowski's use of the locution "Here-Now" show conversely that the special theory of relativity makes essential use of psychological temporal categories in its assertive content (as distinct from the pragmatics of its verification by us humans).

⁵ Cf. H. Bergson: *Creative Evolution* (New York: Random House, Inc.; 1944) and *Matière et Mémoire* (Geneva: A. Skira; 1946). Related criticisms of Bergson's treatment of other aspects of time are given in A. Grünbaum: "Relativity and the Atomicity of Becoming," *op. cit.*, pp. 144-55.

⁶ H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, p. 116.

⁷ The metaphor "crawling" must not, of course, be taken to suggest the "metaphysical error" charged against it by J. J. C. Smart ["Spatializing Time," *Mind*, Vol. CXIV (1955), p. 240] that psychologically time itself "flows" spatially at a certain rate measured in some non-existent hypertime. We shall see shortly that the concept of the "forward" shifting now does not involve this logical blunder.

But this picture of a "block universe," composed of a timeless web of "world-lines" in a four-dimensional space, however strongly suggested by the theory of relativity, is a piece of gratuitous metaphysics. Since the concept of change, of something happening, is an inseparable component of the common-sense concept of time and a necessary component of the scientist's view of reality, it is quite out of the question that theoretical physics should require us to hold the Eleatic view that nothing happens in "the objective world." Here, as so often in the philosophy of science, a useful limitation in the form of representation is mistaken for a deficiency of the universe.⁸

But contrary to Black, Weyl's claim that the time of inanimate nature is devoid of *happening* in the sense of *becoming* is not at all tantamount to the Eleatic doctrine that *change* is an illusion of the human mind! It is of the essence of the relativistic account of the inanimate world as embodied in the Minkowski representation that there is change in the sense that different kinds of events can (do) occur at different times: the attributes and relations of an object associated with any given world-line may be different at different times (e.g., its world-line may intersect with different world-lines at different times). Consequently, the total states of the world (when referred to the simultaneity criterion of a particular Galilean frame) are correspondingly different at different times, i.e., they change with time. It is Black's own misidentification of mere change with becoming ("happening") which leads him to the astonishing and grotesque supposition that Weyl's mentalistic account of becoming bespeaks Weyl's unawareness that "the concept of change . . . is . . . a necessary component of the scientist's view of reality." Black refers to the web of earlier-later relations represented by the world-lines as "timeless" just because they do not make provision for becoming. And he suggests that Weyl conceives of them as forming a four-"space" in the sense in which physical space *excludes* the system of temporal relations obtaining with respect to "earlier than." It is apparent that Black's use of the terms "timeless" and "space" in this context is misleading to the point of conveying question-begging falsehoods. Weyl's thesis is that coming *into* being ("happening"), as con-

⁸ M. Black: "Review of G. J. Whitrow's *The Natural Philosophy of Time*," *Scientific American*, Vol. CCVI (April, 1962), pp. 181-82.

trasted with simply being, is only coming into the present awareness of a sentient organism. And that thesis is not vulnerable to Black's charge of having mistaken "a useful limitation in the form of representation" for "a deficiency of the universe," all the less so, since Weyl makes a point of the difference between space and time by speaking of the world as "a $(3 + 1)$ — dimensional metrical manifold"⁹ rather than as a 4-dimensional one.

In an endeavor to erect a *reductio ad absurdum* of Weyl's thesis, M. Capek has given an even more grotesque account of that thesis than Black did. Capek writes:

although the world scheme of Minkowski eliminates succession in the physical world, it recognizes at least the *movement of our consciousness* to the future. Thus arises an absurd dualism of the timeless physical world and temporal consciousness, that is, a dualism of two altogether disparate realms whose correlation becomes completely unintelligible. . . . in such a view . . . we are already dead without realizing it now; but our consciousness creeping along the world line of its own body will certainly reach any pre-existing and nominally future event which in its completeness *waits* to be finally reached by our awareness. . . . To such strange consequences do both spatialization of time and strict determinism lead.¹

But Capek states a careless and question-begging falsehood by declaring that on Weyl's view the physical world is "timeless." For what Weyl is contending is only that the physical world is devoid of becoming, while fully granting that the states of physical systems are ordered by an "earlier than" relation which is isomorphic, in important respects, with its counterpart in consciousness. Capek's claim of the unintelligibility of the correlation between physical and psychological time within Weyl's framework is therefore untenable, especially in the absence of an articulation of the kind (degree) of correlation which Capek requires and also of a *justification* of that requirement. More unfortunate still is the grievous mishandling of the meaning of Weyl's metaphor in Capek's attempt at a *reductio ad absurdum* of Weyl's view, when Capek speaks of our "already" being dead

⁹ H. Weyl: *Space-Time-Matter*, *op. cit.*, p. 283.

¹ M. Capek: *The Philosophical Impact of Contemporary Physics* (Princeton: D. Van Nostrand Company, Inc.; 1961), p. 165.

without realizing it now and of our completed future death *waiting* to be finally "reached" by our awareness. This gross distortion of Weyl's metaphorical rendition of the thesis that coming *into* being is only *coming into present awareness* rests on a singularly careless abuse of the temporal and/or kinematic components of the meanings of the words "already," "completed," "wait," "reach," etc.

I have argued that in the psychological (common sense) context to which the transient now does have relevance, it is tautological to assert that time "flows" from the past to the future or from earlier to later. But I wish to conclude by noting that it is unjustified to charge the latter assertion with a breach of logical grammar.

We saw that the concept of the transiency of the now or of the flow of time is a qualitative concept without any metrical ingredients. Hence the entirely *non*-metrical concept represented by the metaphor "forward flow" is not at all vulnerable to the metrical *reductio ad absurdum* offered by J. J. C. Smart² in the following form: "The concept of the flow of time or of the advance of consciousness is, however, an illusion. How fast does time flow or consciousness advance? In what units is the rate of flow or advance to be measured? Seconds per — ?" Max Black³ likewise unwarrantedly asks the *metrical* question "How fast does time flow?" and then goes on to claim quite mistakenly that a non-existent super-time would be needed to give meaning to the flow of psychological time. Furthermore Black supposes incorrectly that the metaphor "forward flow" commits one to saying that *time* is "changing"⁴ and to the contention that it makes sense to speak of the cessation of the flow of psychological time. The absurdities which Black is then able to derive from the literal interpretation of these latter assertions can therefore *not* serve to discredit the concept of the transiency of the Now as understood in this book.

² J. J. C. Smart: "The Temporal Asymmetry of the World," *op. cit.*, p. 81.

³ M. Black: "The 'Direction' of Time," *op. cit.*, p. 57.

⁴ *Ibid.*

Chapter II

EMPIRICISM AND THE THREE-DIMENSIONALITY OF SPACE

The success of empiricism in accounting for our knowledge of the tri-dimensionality of the physical world is intimately connected with its ability to refute Kant's claim that the existence of such *similar* but *incongruent* counterparts as the left and right hands constitutes evidence for his transcendental a priori of space.¹ Since the reasons for the untenability of *this particular* Kantian contention are not given even in Reichenbach's definitive empiricist critique of the transcendental idealist theory of space² and are not sufficiently known to the philosophical public, I shall give a brief statement of them.

If we take two arbitrarily (irregularly) shaped objects in a Euclidean plane, which are metrically symmetric or "reflected" about a straight line in that plane, it will be seen that so long as we confine these two objects to that plane, they cannot be brought into congruence such that the points of one coincide with their respective image points in the other. But such congruence *can* be achieved, if we are allowed to rotate one of these reflected two-dimensional objects about the axis of symmetry, thereby making use of the next higher (third) dimension. G. Lechalas credits Delboeuf with discovering that, in

¹ I. Kant: *Werke*, edited by E. Cassirer (Berlin: Bruno Cassirer Verlag; 1912), Vol. II, pp. 393–400 and Vol. IV, §13, pp. 34–36.

² H. Reichenbach: "Kant und die Naturwissenschaft," *Die Naturwissenschaften*, Vol. XXI (1933), pp. 601–606 and 624–26; *Relativitätstheorie und Erkenntnis Apriori* (Berlin: J. Springer; 1920), and *PST*, *op. cit.*

general, given two $(n-1)$ -dimensional objects, metrically symmetric about some $(n-2)$ -dimensional object, then to achieve congruence such that the points of the one coincide with their respective image points in the other, a continuous rotation in n -dimensional space is necessary.³ It should be pointed out, however, that the requirement of rotation through a hyper-space does not hold unrestrictedly: it does hold for spaces whose topology is Euclidean or spherical, but it fails to hold for the Möbius strip two-space or for a one-dimensional space whose topology is that of the numeral 8.

Accordingly, the three-dimensional right-hand cannot be brought into the stated sort of congruence with the three-dimensional left-hand by a continuous rigid motion, because of the *empirical* fact that the four-dimensional space needed for the required kind of rotation is physically unavailable! This same fact enables us to infer the *three*-dimensionality as opposed to the two-dimensionality of optically active molecules from their dextro-rotary or levo-rotary behavior. For if they were only two-dimensional, then it would be possible to convert a given dextro-rotary molecule into a levo-rotary one by merely flipping it over. But this cannot be done.⁴

Contrary to Kant, the specific structural difference between the right and left hands can be given a conceptual rather than *only* a denotatively intuitive characterization as follows:⁵ the group of Euclidean rigid motions is only a *proper* sub-group of the group of length-preserving ("non-enlarging") similarity mappings. For the determinant of the coefficients of the particular linear transformations constituting the latter type of similarity mappings must have either the value +1 or the value -1. But only those similarity transformations whose determinant

³ G. Lechalas: "L'Axiome de libre Mobilité," *Revue de Métaphysique et de Morale*, Vol. VI (1898), p. 754. This property of reflections had already been pointed out by Möbius in his *Der Barycentrische Calcul* (Leipzig: Barth; 1827), p. 184.

⁴ Cf. John Read: *A Direct Entry to Organic Chemistry* (London: Methuen & Company, Ltd.; 1948), Chapter vii.

⁵ Cf. F. Klein: *Elementary Mathematics From An Advanced Standpoint*, (New York: Dover Publications, Inc.; 1939), Vol. II, pp. 39-42; H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, pp. 79-85. For a more elementary account, see O. Hölder: *Die Mathematische Methode*, *op. cit.*, pp. 387-89.

("Jacobian")⁶ is +1 form the group of Euclidean rigid motions, the remainder being the *reflections* whose Jacobian is -1 and which *include* the case of *Kant's left and right hands*.⁷ But the very existence of the latter pairs of physical objects in our physical three-space which realize a formal transformation whose Jacobian is -1 is an *empirical* fact, as is the *non-existence* of a *physical* four-dimensional hyper-space through which they could otherwise be rotated so as to be brought into congruence.

See Append. §32

Since the tri-dimensionality of physical space has turned out to be a logically contingent empirical fact, one naturally wonders whether it is an autonomous, irreducible empirical fact or not with respect to the standard formulations of physical theory. Huyghens's principle in optics tells us that if a single spherical light wave is produced by a disturbance at a point which lasts for a very short time between $t = t_0 - \epsilon$ and $t = t_0$, then the effect at a point P at a distance cT (where c = the velocity of light) is null until the instant $t = t_0 - \epsilon + T$ and is null again after the instant $t = t_0 + T$. And thus, according to Huyghens's principle, a single spherical wave would leave no residual after-effect at a point P. Now, J. Hadamard has shown⁸ that this requirement of Huyghens's principle is satisfied only by wave equations having an *even* number of independent variables. Since the time variable in conjunction with the space variables constitute the independent variables of these equations, Hadamard's result

⁶ Let F_1 and F_2 be two functions of u and v . Then

$$J = \begin{vmatrix} \frac{\partial F_1}{\partial u} & \frac{\partial F_1}{\partial v} \\ \frac{\partial F_2}{\partial u} & \frac{\partial F_2}{\partial v} \end{vmatrix} = \frac{\partial F_1}{\partial u} \cdot \frac{\partial F_2}{\partial v} - \frac{\partial F_1}{\partial v} \cdot \frac{\partial F_2}{\partial u} \text{ is the functional}$$

determinant or "*Jacobian*" of F_1 and F_2 with respect to u and v . In the three-dimensional case relevant to the incongruent hands, we are dealing with three functions F_1 , F_2 , and F_3 of u , v , and w , and the corresponding Jacobian is a determinant of three rows and three columns.

⁷ A requirement for the preservation of the dimensionality under a transformation is that the Jacobian *not* be equal to zero, the latter condition being necessary and sufficient that the transformation be one-to-one. Cf R. S. Burington and C. C. Torrance: *Higher Mathematics* (New York: McGraw-Hill Book Company, Inc.; 1939), pp. 132-42.

⁸ J. Hadamard: *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (New Haven: Yale University Press; 1923), pp. 53-54, 175-77, and 235-36.

shows that Huyghens's principle holds only for cases in which the number of *space* dimensions is *odd*, as in the case of the three-dimensional physical space of our world.⁹ Sharing the view of Aristotle and Galileo that the tri-dimensionality of physical space might be explainable as a consequence of other, more comprehensive empirical principles, H. Weyl suggests¹ that the difference between spaces of even and odd numbers of dimensions in regard to the transmission of waves may be one clue to the required explanation.²

Contrary to a widely held interpretation, Poincaré's account of the status of the tri-dimensionality of space is not incompatible with the empirico-realistic conception of that attribute of the physical world just set forth. In his rejoinder to Russell, we find Poincaré saying without elaboration: "I consider the axiom of three dimensions as conventional in the same way as those of Euclid."³ But in his posthumous book,⁴ he tells us that since his earlier treatment of this axiom was "very compressed," he now wishes to clarify it. Then, after explaining that in classifying the elements of a manifold as the same in some respect, we use the "basic convention" of abstracting from other qualitative differences among them, he notes that the three-dimensionality of the perceptual localizations of physical events is obtained upon abstracting from a variety of qualitative *non-positional* differences between them. This sense of "convention," however, hardly renders three-dimensionality non-objective any more than the reference to kinds of events makes particular causal statements true by convention. That Poincaré was entirely clear on this is apparent from the following assertion by

⁹ For an explanation of this result by reference to the special case of three and two dimensions, cf. B. Baker and E. T. Copson: *The Mathematical Theory of Huyghens's Principle* (Oxford: Oxford University Press; 1939), pp. 46–47.

¹ H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, p. 136.

² Cf. also P. Ehrenfest: "In What Way Does It Become Manifest in the Fundamental Laws of Physics That Space Has Three Dimensions?" *Proceedings of the Amsterdam Academy*, Vol. XX (1917), pp. 200–209, or, in German translation: "Welche Rolle spielt die Dreidimensionalität des Raumes in den Grundgesetzen der Physik," *Annalen der Physik*, Vol. LXI (1920), pp. 440–46.

³ H. Poincaré: "Sur les Principes de la Géométrie," *op. cit.*, p. 73.

⁴ H. Poincaré: *Letzte Gedanken*, *op. cit.*, p. 59.

him: "We see on the basis of this brief explanation what experimental facts lead us to ascribe three dimensions to space. As a consequence of these facts, it would be more convenient for us to attribute three dimensions to it than four or two; but the term 'convenient' is perhaps not strong enough; a being which had attributed two or four dimensions to space would be handicapped in a world like ours in the struggle for existence."⁵ After exhibiting how that handicap would arise from an interpretation of space as two- or four-dimensional, he shows on the basis of group-theoretical arguments⁶ that in the context of causality, physical facts lead to the tri-dimensionality of physical space just as the structure of perceptual data had. And he concludes by saying that since we have the capacity to construct mathematically a continuum of an arbitrary number of dimensions, "this capacity would . . . permit us to construct a space of four dimensions just as well as one of [only] three dimensions. *It is the external world, experience*, which determines our developing our ideas more in one of these directions than in the other."⁷

The mathematical continuity which is attributed to physical space is a *topological* property just as its tri-dimensionality. We shall therefore now conclude our consideration of philosophical problems of the *topology* of time and space in Part II by inquiring whether continuity can be held to have any kind of *empirical* status.

We argued in Chapter One that the continuity postulated for physical space and time issues in the metric amorphousness of these manifolds and thus makes for the conventionality of congruence, much as the conventionality of non-local metrical simultaneity in special relativity is a consequence of the postulational *fact* that light is the fastest causal chain *in vacuo* and that clocks do not define an absolute metrical simultaneity under transport. But it has been objected that there is an important epistemological difference between the latter Einstein postulate and the postulational ascription of continuity (in the mathematical sense) to physical space and time instead of some discontinuous structure: the continuity postulate, it is held, cannot be regarded,

⁵ *Ibid.*, p. 86.

⁶ *Ibid.*, Chapter iii, §5, pp. 87-94.

⁷ *Ibid.*, p. 99, my italics. Cf. also O. Hölder: *Die Mathematische Methode*, *op. cit.*, p. 393.

even in principle, as a *factual* assertion in the sense of being either true or false. For according to this objection, there can be no *empirical* grounds for accepting a geometry postulating continuous intervals in preference to one which postulates discontinuous intervals consisting of, say, only the algebraic or only the rational points. The objector therefore contends that the rejection of the latter kind of denumerable geometry in favor of a non-denumerable one affirming continuity has no kind of factual warrant but is based solely on considerations of *arithmetic convenience* within the analytic part of geometry. And hence the *topology* is claimed to be no less infested with features springing from conventional choice than is the metric. Accordingly, the exponent of this criticism concludes that it is not only a misleading emphasis but outright incorrect for us to discern conventional geometrical elements only in the *metrization* of the topological substratum while deeming the *continuity* of that substratum to be *factual*.

This plea for a conventionalist conception of continuity is not convincing, however. Admittedly, the justification for regarding continuity as a broadly *inductive framework-principle* of physical geometry cannot be found in the direct verdicts of measuring rods, which could hardly disclose the super-denumerability of the points on the line. And *prima facie* there is a good deal of plausibility in the contention that the postulation of a super-denumerable infinity of irrational points in addition to a denumerable set of rational ones is dictated solely by the desire for such arithmetical convenience as having closure under operations like taking the square root, etc. But, even disregarding the Zenonian difficulties which may vitiate denumerable geometries logically, as we saw in Chapter Six, these considerations lose much of their force, it would seem, as soon as one applies the *acid test* of a convention to the conventionalist conception of continuity in physical geometry: the availability or demonstrated feasibility of one or more *alternate* formulations dispensing with the particular alleged convention and yet permitting the successful theoretical rendition of the same total body of experimental findings, such as in the case of the choice of a particular system of *units* of measurement.

Upon applying this test, what do we find? No mathematically discontinuous alternative set of theories have been shown to be

as viable as those which are based on the continuum by demonstrating that these two kinds of theories must be, *in principle*, empirically indistinguishable from one another.⁸ In the absence of such a demonstration, let alone of the actual, full-fledged elaboration of an empirically adequate discontinuous alternative theory, the charge that in a geometry the topological component of continuity is no less conventional than the metric itself is *unfounded*. And pending the elaboration of a successful alternative to the continuum, the following suspicion insinuates itself: the empirical facts codified in terms of the classical mathematical apparatus in our most sophisticated and best-confirmed physical theories support continuity in a *broadly inductive* sense as a *framework principle* to the exclusion of the *prima facie* rivals of the continuum.

It would be an error to believe that this conclusion requires serious qualification in the light of recent suggestions of space (or time) quantization. For as H. Weyl has noted

so far it [i.e., the atomistic theory of space] has always remained mere speculation and has never achieved sufficient contact with reality. How should one understand the metric relations in space on the basis of this idea? If a square is built up of miniature tiles, then there are as many tiles along the diagonal as there are along the side; thus the diagonal should be equal in length to the side.⁹

And Einstein has remarked that:

From the quantum phenomena it appears to follow with certainty that a finite system of finite energy can be completely described by a finite set of numbers (quantum numbers). This does not seem to be in accordance with a continuum theory, and must lead to an attempt to find a purely algebraic theory for the description of reality. But nobody knows how to obtain the basis of such a theory.¹

⁸ P. K. Feyerabend has suggested ["Comments on Grünbaum's 'Law and Convention in Physical Theory,'" in: H. Feigl and G. Maxwell (eds.) *Current Issues in the Philosophy of Science*, *op. cit.*, pp. 160-61] that a crucial experiment between continuous and discontinuous descriptions of nature is logically possible by giving a sketch of such an experiment.

⁹ H. Weyl: *Philosophy of Mathematics and Natural Science*, *op. cit.*, p. 43.

¹ A. Einstein: *The Meaning of Relativity* (Princeton: Princeton University Press; 1955), pp. 165-66.

By claiming to have established the *unfoundedness* of the conventionalist conception of continuity, I do not, of course, maintain that I have demonstrated its falsity. P. K. Feyerabend has asserted² that the absence at this time of a discontinuous alternative to current physical theory "does not prove anything," since "after all, the full theory of continuity was not developed before the second half of the nineteenth century." But Feyerabend overlooked here that my argument does establish the *gratuitousness* of the conventionalist conception of continuity by showing that, at best, its advocates are offering only a *program*.

Can my argument be strengthened by the failure of neo-intuitionist attempts to dispense with the continuum of geometry and analysis via the provision of substitutes for all of classical mathematics? These neo-intuitionistic endeavors to base classical mathematics on more restrictive foundations failed by involving truncations of mathematical physics whose range A. A. Fraenkel has characterized as follows:

intuitionistic restriction of the concept of continuum and of its handling in analysis and geometry, though carried out in quite different ways by various intuitionistic schools, always goes as far as to exclude vital parts of those two domains. (This is not altered by Brouwer's peculiar way of admitting the continuum *per se* as a "medium of free growth.")³

But, as Feyerabend has pointed out correctly, the failures of the neo-intuitionists cannot be regarded as strengthening my argument:

For what we want is not repetition, in new terms, of the *whole* of classical mathematics; what we want is a new mathematical system that may be adequate for the description of the universe which is also articulate enough for a crucial experiment to be possible.⁴

² P. K. Feyerabend: "Comments on Grünbaum's 'Law and Convention in Physical Theory,'" *op. cit.*, p. 160.

³ A. A. Fraenkel and Y. Bar-Hillel: *Foundations of Set Theory* (Amsterdam: North Holland Publishing Company; 1958), p. 200. Chapter iv of this work gives an admirably comprehensive and lucid survey of the respects in which neo-intuitionist restrictions involve mutilations of the system of classical mathematics.

⁴ P. K. Feyerabend: "Comments on Grünbaum's 'Law and Convention in Physical Theory,'" *op. cit.*, p. 161.

PART III

*Philosophical Issues
in the Theory
of Relativity*

Chapter 12

PHILOSOPHICAL FOUNDATIONS OF THE SPECIAL THEORY OF RELATIVITY, AND THEIR BEARING ON ITS HISTORY

(A) INTRODUCTION

Since the publication of Hans Reichenbach's definitive books on the philosophy of the theory of relativity during the nineteen twenties,¹ the literature on the philosophy and history of Einstein's special theory of relativity (hereafter called STR) has been enlarged by contributions which call for critical evaluation. In this chapter, I shall give such an evaluation in the course of presenting: (a) an up-to-date analysis of the intertwined philosophical and empirical foundations of the kinematics of the STR, with attention to neglected issues and prevalent misconceptions, and (b) a demonstration that a rigorous grasp of the philosophical conceptions underlying the fully evolved STR and distinguishing it from its ancestors is decisively prerequisite to (1) the very posing of well-conceived, searching historical questions in regard to the STR, and hence to (2) the provision of a historically sound and illuminating account of its genesis.

Specifically, it will be useful to interlace the presentations of (a) and (b) in sections as follows:

- (B) Einstein's conception of simultaneity, its prevalent misrepresentations, and its history,

¹ H. Reichenbach: *Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, Vol. 72 of *Die Wissenschaft*, and *PST*, *op. cit.*

- (C) The history of Einstein's enunciation of the limiting character of the velocity of light *in vacuo*,
- (D) The principle of the constancy of the speed of light, and the falsity of the aether-theoretic Lorentz-Fitzgerald contraction hypothesis,
- (E) The experimental confirmation of the kinematics of the STR, and
- (F) The philosophical issue between Einstein and his aether-theoretic precursors; its bearing on E. T. Whittaker's history of the STR.

Although the historical portions of this chapter will emphasize the indispensability of philosophical mastery of the STR to the unraveling of its history, I do not wish to deny that the history of the STR, in its turn, can be relevant to the understanding of the philosophical ingredients of that theory. On the contrary, historical inquiry may be illuminating philosophically by disclosing, for example, the vicissitudes in Einstein's own philosophical orientation, thereby explaining, as G. Holton has pointed out, why advocates of contending schools of philosophic thought can each "find some part of Einstein's work to nail to his mast as a battle flag against the others."² Despite the capability of the history of a theory to serve as a propaedeutic to the analysis of its philosophical foundations, it happens that what is known reliably so far about the history of the STR has failed signally, as we shall see, to contribute to the clarification of the more subtle questions concerning its *epistemological* basis. And it will become apparent that some of the historical accounts which have been given are beset by serious and puzzling contradictions.

(B) EINSTEIN'S CONCEPTION OF SIMULTANEITY, ITS PREVALENT MISREPRESENTATIONS, AND ITS HISTORY.

Einstein discusses the problem of simultaneity epistemologically in Section 1 of his 1905 paper³ well before showing at the

² G. Holton: "On the Origins of the Special Theory of Relativity," *American Journal of Physics*, Vol. XXVIII (1960), p. 627.

³ A. Einstein: "On the Electrodynamics of Moving Bodies," *op. cit.*, pp. 38-40. The cited collection containing this paper will be cited hereafter as "PR."

end of Section 2 that the two fundamental postulates of his STR entail discordant judgments of simultaneity as between relatively moving Galilean frames. And he stresses in Section 1, though only rather concisely, that, contrary to Newtonian theory, the metrical simultaneity of two spatially separated events involves a convention *within* any given inertial system. Nevertheless, numerous expositions of the STR have nurtured the misunderstanding that the philosophical repudiation of Newton's absolute simultaneity by the STR occurs, *in the first instance*, in the context of the theory of the relative motion of different inertial systems. We shall see in detail why it is an error to regard the discordant judgments of simultaneity made by different Galilean frames as the *primary* locus of the conceptual innovation wrought by Einstein in regard to the concept of simultaneity. This misunderstanding has the following consequences: First, it obscures the fact that Einstein's philosophical supplanting of Newton's conception of simultaneity is *presupposed* by rather than first derived from his enunciation of the fundamental postulate of the constancy of the speed of light, and second, it precludes awareness that Einstein's philosophical characterization of distant simultaneity as *conventional* rests on specifiable *physical assumptions*, thereby preventing the recognition of the logical role played by these particular physical assumptions in the very foundations of the STR.

The important bearing of the philosophical misconstrual of Einstein's conception of simultaneity on the investigation of the history of the STR is therefore the following: the lack of philosophical mastery on the part of the historian will conceal from him that there exists the historical problem as to the grounds on which Einstein felt entitled in 1905 to make the particular physical assumptions undergirding his philosophical doctrine of the conventionality of the simultaneity of spatially separated events.

This doctrine is set forth *very concisely* in Section 1 of his 1905 paper, which is entitled "Definition of Simultaneity." There he writes: "We have not defined a common 'time' for [the spatially separated points] A and B, for the latter cannot be defined at all unless we establish *by definition* that the 'time' required by light to travel from A to B equals the 'time' it re-

quires to travel from B to A."⁴ This conception of simultaneity rests on a number of philosophical and physical assumptions which we shall now discuss.

It is clear that the sense in which two spatially separated events can be simultaneous is different from the one in which two quasi-coinciding events can be regarded as simultaneous: in the latter case of quasi-coincidence, local or contiguous simultaneity obtains, and not simultaneity at a distance. It might be thought that distant simultaneity can be readily grounded on local simultaneity as follows. Suppose that two spatially separated events E_1 and E_2 produce effects which intersect at a sentient observer so as to produce the experience of sensed (intuitive) simultaneity at himself. Then this local simultaneity of the effects of E_1 and E_2 permits us to infer that the distant events occurred simultaneously, if the influence chains that emanated from them had appropriate *one-way velocities* as they traversed their respective distances to the location of the sentient observer. But this procedure is unavailing for the purpose of first characterizing the conditions under which the separated events E_1 and E_2 can be held to be simultaneous. For the *one-way velocities* invoked by this procedure presuppose *one-way transit times* which are furnished by *synchronized* clocks at the locations of E_1 and E_2 respectively. And the conditions for the synchronism of two spatially separated clocks, in turn, presuppose a criterion for the distant simultaneity of the events at these clocks.⁵

These considerations enable us to demonstrate the complete failure of A. N. Whitehead's attempt to ground the concept of distant simultaneity of physical theory on the sensed coincidences experienced by sentient observers.⁶ In the first place, there are inconsistencies between the sensed simultaneity verdicts concerning a given pair of separated events when furnished

⁴ A. Einstein: *PR*, p. 40; *italics in the original*. But, as has been correctly noted by C. Scribner ("Mistranslation of a Passage in Einstein's Original Paper on Relativity," *American Journal of Physics* [1963], Vol. XXXI, p. 398), the middle part of this sentence was mistranslated into English by Perrett and Jeffery and should read instead as follows: "the latter time can now be defined in establishing *by definition* that the 'time,' etc."

⁵ H. Reichenbach: *Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, pp. 12-17, and *PST*, *op. cit.*, pp. 123-26.

⁶ A. N. Whitehead: *The Concept of Nature*, *op. cit.*, pp. 53 and 56.

by sentient observers who are stationed at different space points of the same inertial system: if influences emanating from two spatially separated events E_1 and E_2 intersect at a point P_1 so as to produce sensed coincidence in the awareness of a sentient observer stationed at P_1 , then there will be *other* points P_2 , P_3 , . . . , P_n in the *same* inertial system at which these same events E_1 and E_2 will *not* produce sensed coincidence in the respective stationary sentient observers.⁷ And if Whitehead is to avoid these *inconsistencies* between the diverse verdicts as to sensed simultaneity in a manner coherent with physical theory, he must *restrict* the ascription of simultaneity on the basis of sensed simultaneity to only the following kinds of separated events: Those whose simultaneity would be compatible with the physical one-way velocities of the influence chains emanating from these events and meeting at the sentient observer after traversing the respective given distances. But the information concerning one-way velocities which Whitehead requires in order to be safe in interpreting the sensed coincidence as signifying simultaneity presupposes a prior criterion of simultaneity or clock-synchronization, which sensed coincidence is manifestly incompetent to furnish without involving Whitehead in a vicious circle.

One must therefore endorse Einstein's rejection of sensed simultaneity as the basis for distant physical simultaneity. His constructive alternative to Whitehead's unsuccessful attempt can be understood on the basis of his repudiation of both philosophical and physical assumptions made by the Newtonian theory as follows.

In opposition to the absolutistic conception of space and time ingredient in the Newtonian theory, the STR rests on the following conceptions of the identity of events and of their temporal order: first, prior to the construction of any system of coordinates, physical things and events first define, by their own identity, points and instants which can then constitute the spatial and temporal loci of other physical objects and events; and second, the temporal order of physical events is first constituted by their properties and relations qua physical events. By maintaining that the very *existence* of *temporal* relations between non-coinciding

⁷ Cf. F. S. C. Northrop: "Whitehead's Philosophy of Science," *op. cit.*, p. 200.

events depends on the obtaining of some *physical* relations between them, Einstein espoused a conception of time (and space) which is *relational* by regarding them as systems of relations between physical events and things. Since time relations are first constituted by the system of physical relations obtaining among events, the character of the temporal order will be determined by the physical attributes in virtue of which events will be held to sustain relations of "simultaneous with," "earlier than," or "later than." In particular, it is a question of physical fact whether these attributes are of the kind to define temporal relations *uniquely*—i.e., such that on the strength of the obtaining of these attributes, *every* pair of events is *unambiguously* ordered with respect to one of the relations "earlier than," "later than," or "simultaneous with." If the temporal order were to have the latter character, it would be "absolute" in the compound sense that the time relation between any two events would be a *uniquely* obtaining *factual* relation between them which is wholly independent of any particular reference system and hence the same in every reference system. Were it to obtain in the case of simultaneity, it would have the consequence that *as a matter of physical fact, one and only one event* at a point Q would be simultaneous with a given event occurring at a point P elsewhere in space, a state of affairs which I have characterized by saying for brevity that simultaneity would be a "*uniquely* obtaining *factual* relation."

But is there in actual fact a physical basis for relations of *absolute* simultaneity among spatially separated events? If the behavior of transported clocks were of the kind assumed by the Newtonian theory, then indeed absolute simultaneity could be said to obtain merely in virtue of the coincidence of physical events with suitable readings of transported clocks of identical constitution. But the STR denies the existence of precisely this physical-clock basis for absolute simultaneity. For it makes the following assumption, to which we shall refer as "assumption (i)": within the class of physical events, material clocks do not define relations of absolute simultaneity under transport, because the relations of simultaneity which are defined by transported clocks *do* depend on the particular clock that is used, in the following sense: If two clocks U_1 and U_2 are initially synchro-

nized at the same place A and then transported via paths of *different lengths* to a different place B such that their arrivals at B coincide, then U_1 and U_2 will no longer be synchronized at B. And if U_1 and U_2 were brought to B via the same path (or via different paths which are of equal length) such that their arrivals do *not* coincide, then too their initial synchronization would be destroyed. Thus, depending upon which one of the discordant clocks would serve as the standard, a given pair of events at A and B *would* or *would not* be held to be simultaneous. Hence, this dependence on the particular clock used prevents transported clocks from defining relations of absolute simultaneity within the class of physical events. And it *also* prompted Einstein to *deny* that even *within* a single inertial system, the physical basis of the simultaneity of two spatially separated events E and E° can be constituted, in part, by the readings produced by clocks as follows: two clocks U_1 and U_2 are transported to separate places from a common point in space at which they had identical readings, and then event E occurs at the location of clock U_1 , and event E° at the place of clock U_2 such that the readings of U_1 and U_2 are the same.

The failure of transported clocks to constitute relations of absolute simultaneity among spatially separated events does not, of course, show that there is also no other physical relatedness among events which would make for such relations. But non-coinciding and *spatially separated* events can sustain physical relations of one kind or another only in virtue of the presence or absence of some kind of actual (or physically possible) physical linkage between them. And it would seem that the only direct linkage connecting the latter kinds of events can be constituted by *causal chains* of which they are the termini.

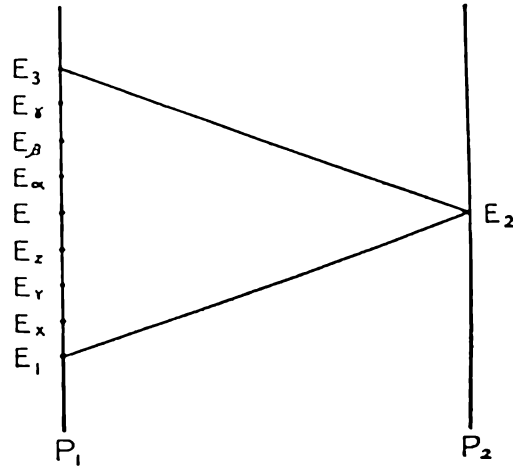
Let us consider, therefore, under what conditions a pair of spatially separated events can be simultaneous, if the relations of temporal order between such events depend on the obtaining or non-obtaining of a physical relation of *causal connectibility* between them. But we recall from Chapters Seven and Eight that (1) the characterization of *one* of two causally connected (or connectible) and genidentical events as "the" (possible partial or total) cause of the other presupposed the anisotropy of time, and (2) our account of the anisotropy of time, i.e., of the

physical basis of the relations "earlier than" and "later than," in turn, presupposed distant simultaneity. Hence our impending statement of the bearing of the non-obtaining of causal connectibility on simultaneity will invoke only the symmetric causal relation of Chapter Seven, which is non-committal as to whether there are physical criteria for singling out one of two causally connected events as *the* (partial or total) cause of the other. And only after the concept of simultaneity ingredient in the account of the anisotropy of time thus becomes logically available will we invoke the latter to characterize one of two causally connected events as the earlier of the two.

Accordingly, we now assert: if and only if two *non-coinciding* events sustain the symmetric relation of genidentical (or quasi-genidentical) causal *connectibility* discussed in Chapter Seven will these two events be said to sustain the relation of *temporal separation*, i.e., the relation of being either earlier or later. And, therefore, two events will be said to sustain the relation of being neither earlier nor later, i.e., the relation of being *topologically simultaneous* or of having a *spacelike* separation, if and only if the specified kind of causal connectibility does *not* obtain between them.

If the physical basis for the relation of topological simultaneity among events lies in the impossibility of their being the termini of influence chains, we ask: Are the physically possible causal chains of nature such as to define a *unique* set of temporal relations in which every pair of events has an unambiguous place with respect to the relations of temporal separation and topological simultaneity? To answer this question, consider four events E_1 , E_2 , E_3 , and E which satisfy the following conditions: First, E_1 , E_2 , and E_3 are genidentically connectible *by a light ray* such that E_2 is temporally o-between E_1 and E_3 , and second, E_1 , E , and E_3 are genidentically connectible other than by a light ray such that E is temporally o-between E_1 and E_3 , and all three of these events occur at the *same* space point P_1 of an inertial system, while E_2 occurs at a different space point P_2 of that same inertial system. Furthermore, let E_x , E_y , and E_z be events temporally o-between E_1 and E such that E_x is o-between E_1 and E_y , while E_y is o-between E_x and E_z . And finally, let E_a ,

E_β , and E_γ be o-between E and E_3 as shown in the following diagram, which is a world-line diagram but devoid of any of the arrows that might invoke the anisotropy of time as a basis for an (intrinsic) characterization of the positive time direction. As will be recalled from Chapters Seven and Eight, the latter direction could, of course, be introduced purely extrinsically by a mere coordinatization as effected through standard clocks. But we shall postpone the introduction of such a coordinatization, since it will be instructive to see that Einstein's thesis of the conventionality of metrical simultaneity can be formulated without it.



It is clear that the given relations that E_2 is temporally o-between E_1 and E_3 , and that E , E_x , E_y , E_z , E_α , E_β , and E_γ are each o-between E_1 and E_3 as specified do not furnish any basis for relations of temporal separation or topological simultaneity between E_2 and any one of the events at P_1 lying within the open interval between E_1 and E_3 . The existence of the latter relations will therefore depend on whether or not it is physically possible that there be direct causal chains which could link E_2 to a particular event in the open interval at P_1 .

Though admitting that no light ray or other electromagnetic causal chains could provide a link between E_2 and E_z , or E_2 and E_α , Newtonian physics adduces its second law of motion to assert the physical possibility of other causal chains (e.g., moving par-

See Append. § 17 for
comment on the
super-light causal
chains of
"metarelativity"

ticles) which would indeed do so.⁸ But precisely this latter possibility is *denied* by Einstein in the STR, since he affirms the following topological postulate to which we shall refer as "assumption (ii)"⁹: it is physically impossible that *in vacuo*¹ there be genidentical (or other) causal chains which would link E_2 to any two events one or both of which lies *within* the open time interval $\overline{E_1 E_3}$ such that E_2 would be temporally o-between the two events in question.

Einstein's topological affirmation in assumption (ii) of the limiting role of light propagation *in vacuo* within the class of causal chains has the following fundamental consequence: each one of the entire super-denumerable infinity of events at P_1 within the open time interval between E_1 and E_3 rather than only a single such event is topologically simultaneous with E_2 , topological simultaneity therefore not being a uniquely obtaining relation. For, on Einstein's assumption (ii), none of these events at P_1 can be said to sustain the relation of temporal separation to E_2 (i.e., be either earlier or later than E_2) *as a matter of physical fact*. But to say that *none* of these events at P_1 is physically either earlier or later than E_2 is to say that no one of them is objectively any more entitled to be regarded as metrically simultaneous with E_2 than is any of the others. It is therefore only by *convention* or definition that some *one* of these events comes to be metrically simultaneous with E_2 , the remainder of them in the open interval $\overline{E_1 E_3}$ at P_1 then becoming *conventionally* either earlier or later than E_2 .

⁸ In the context of the Newtonian assumption that transported clocks can furnish a coordinatization of time by serving as indicators of absolute simultaneity, the relevant content of the second law of motion can be stated as follows: no matter what the velocity, the ratio of the force to the acceleration is constant, and particles can therefore be brought to arbitrarily high speeds in excess of the speed c of light by appropriately large forces acting for a sufficient time.

⁹ A. Einstein: "Autobiographical Notes," in: P. A. Schilpp (ed.) *Albert Einstein: Philosopher-Scientist* (New York: Tudor Publishing Company; 1949), p. 53.

¹ An account of the reason for the qualification *in vacuo* can be found in I. E. Tamm: "Radiation of Particles with Speeds Greater than that of Light," *American Scientist*, Vol. XLVII (1959), p. 169, and "General Characteristics of Vavilov-Cherenkov Radiation," *Science*, Vol. CXXXI (1960), p. 206.

Unlike the Newtonian situation, in which there was only a single event E which could be significantly held to be simultaneous with E , the physical facts postulated by relativity require the introduction, *within a single inertial frame* S , of a convention stipulating which *particular* pair of topologically simultaneous events at P_1 and P_2 will be chosen to be metrically simultaneous. We see incidentally that *topological* simultaneity or having a *spacelike* separation is *not* a transitive relation: while E_x is topologically simultaneous with E_2 , and E_2 is, in turn, topologically simultaneous with E , the events E_x and E are *not* topologically simultaneous but temporally separated. To this it has been objected that the "ordinary usage" of the term "simultaneous" in common sense discourse is such as to entail the transitivity of the simultaneity relation.

My reply to this objection is that it is no more incumbent upon Einstein to use the technical theoretical term "topologically simultaneous" so as to name a transitive relation on the strength of the "ordinary usage" of "simultaneous" than it is obligatory in physics to use such terms as "temperature" and "work" in their ordinary sense rather than as homonyms of their "ordinary language" counterparts. For the ordinary usage of temporal (and spatial) terms reflects the conceptual commitments of "ordinary users." And these commitments are frequently circumscribed by scientific and philosophical ignorance. In particular, the present-day advocacy of the imposition of ordinary language restrictions on the term "simultaneous" when used in its topological (i.e., *non-metrical*) sense constitutes an insistence on the retention of Newtonian beliefs and can therefore serve as an impediment to the exposition and understanding of the theory of relativity. Small wonder then that within the philosophy of science, ordinary language analysis has been almost entirely a pernicious obscurantism in so far as it has found adherents. And to the extent that ordinary language analysis has been claimed to be the sole task of philosophy, it has most unfortunately served to convince some scientists that no scientifically relevant conceptual clarification can be expected from any professional philosophers. For the bulk of the practitioners of ordinary language analysis themselves, who have enthroned the "ordinary man" to be the

intellectual arbiter of the world, this state of affairs has provided the comforting assurance that scientific ignorance is not a handicap for a philosopher.

In the light of our foregoing considerations, Einstein's conceptual innovation regarding simultaneity can therefore be summarized somewhat as follows: The time relations among events having been assumed as first constituted by physical relations obtaining between them, these physical relations turned out to be such that topological simultaneity is not a uniquely obtaining relation and hence cannot serve, as it stands, as a metrical synchronization rule for clocks at the spatially separated points P_1 and P_2 . Metrical simultaneity having thus been left indeterminate by both topological simultaneity and by the relativistic behavior of transported clocks, a supplementary *convention* and not merely relevant physical facts must be invoked to assert that an event at P_2 sustains a uniquely obtaining equality relation of metrical simultaneity to an event at P_1 . In short, Einstein's physical innovation is that the physical relatedness which makes for the very existence of the temporal order has the kind of gaps that issue in the non-existence of absolute simultaneity, ascriptions of temporal order *within* these gaps therefore being conventional as follows: they rest on a conventional choice of a *unique* pair of events at P_1 and at P_2 as metrically simultaneous from within the class of pairs of events at P_1 and P_2 which are topologically simultaneous. It is evident that the failure of human measuring operations to disclose relations of absolute simultaneity is therefore only the epistemic consequence of the primary non-existence of these relations.

Before giving mathematical expression to the conventionality of metrical simultaneity, I wish to note that we formulated the conventionality of metrical simultaneity without any reference whatever to the relative motion of different Galilean frames and hence without any commitment whatever to there being a *discordance* or *relativity* of simultaneity as between such frames. And we shall see below in concrete detail that the repudiation of Newtonian simultaneity by the thesis of the conventionality of simultaneity does not depend at all for its validity upon there being disagreement among different Galilean observers as to the simultaneity of pairs of non-coinciding events. On the contrary,

it will become apparent that the conventionality of simultaneity provides the logical framework within which the *relativity* of simultaneity can first be understood: if *each* Galilean observer adopts the particular metrical synchronization rule adopted by Einstein in Section 1 of his fundamental paper² and if the spatial separation of P_1 and P_2 has a component along the line of the relative motion of the Galilean frames, then that relative motion issues in their choosing as metrically simultaneous *different pairs of events* from within the class of topologically simultaneous events at P_1 and P_2 , a result embodied in the familiar Minkowski diagram.

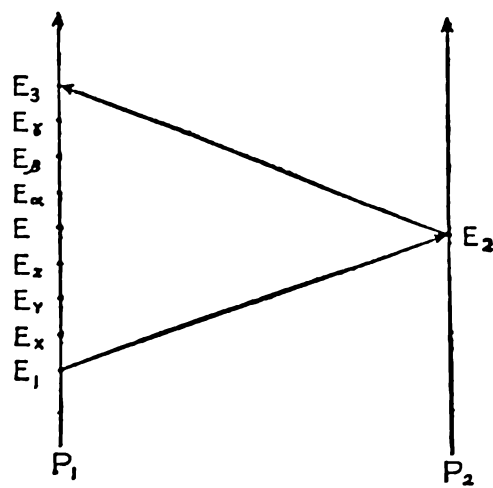
Now let the events at P_1 be assigned time numbers by means of a clock stationed there as follows: t_1 is the time of E_1 , t_3 the time of E_3 and $\frac{1}{2}(t_3 + t_1)$ the time of E . The conventionality of metrical simultaneity then expresses itself in the fact that even *within* the given inertial system S , the obtaining of metrical simultaneity in the system depends on a choice—not dictated by any facts—of a particular numerical value between t_1 and t_3 as the temporal name to be assigned to E_2 at P_2 by an appropriate setting of a like clock stationed there in S . Using Reichenbach's notation,³ we can say, therefore, that depending on the particular event at P_1 that is chosen to be simultaneous with E_2 , upon the occurrence of E_2 we set the clock at P_2 to read

$$(1) \quad t_2 = t_1 + \epsilon (t_3 - t_1),$$

where ϵ has the particular value between 0 and 1 appropriate to the choice we have made. If, for example, we choose ϵ such that E_y —in our now arrow-equipped world-line diagram below—becomes simultaneous with E_2 , then all of the events between E_1 and E_y become definitionally (i.e., not objectively) earlier than E_2 , while all of the events between E_y and E_3 become definitionally later than E_2 . Clearly, we could alternatively choose ϵ such that instead of becoming simultaneous with E_2 , E_y would become earlier than E_2 or later in the same frame S . This freedom to decree definitionally the relations of temporal sequence merely expresses the objective indeterminateness of unique time relations between *causally non-connectible* events; and, of course, such

² A. Einstein: *PR*, *op. cit.*, p. 40.

³ Cf. H. Reichenbach: *PST*, *op. cit.*, p. 127.



WORLD-LINE DIAGRAM

freedom can be exercised only with respect to such pairs of events.⁴ For, as we saw, in the case of causally connectible events the relation of temporal separation has an objective physical basis. And once a criterion of metrical simultaneity has been chosen as outlined, the ensuing time coordinatization makes it unambiguous which one of two objectively time-separated events is the earlier of the two, and which one the later.

It is apparent that no fact of nature found in the objective temporal relations of physical events precludes our choosing a value of ϵ between 0 and 1 which *differs* from $1/2$. But it is only for the value $1/2$ that the velocity of light in the direction P_1P_2 becomes equal to the velocity in the opposite direction P_2P_1 in virtue of the then resulting equality of the respective transit times $t_2 - t_1$ and $t_3 - t_2$. A choice of $\epsilon = 3/4$, for instance, would make the velocity in the direction P_1P_2 only one-third as great as the return velocity. More generally, since the total round-trip time $t_3 - t_1$ is split up into the outgoing one-way time $t_2 - t_1$

⁴ In his posthumously published work entitled *A Sophisticate's Primer of the Theory of Relativity* (Middletown, Connecticut: Wesleyan University Press; 1962 [Chapter ii]), P. W. Bridgman has maintained that there is no need to ground the theory of time order of the STR on causal connectibility and non-connectibility, since time order can allegedly be based in the STR on philosophically self-contained alternative physical foundations. But in my Epilogue to this book by Bridgman, I have argued (*ibid.*, pp. 181-90) that Bridgman's contention is unsound.

and the return one-way time $t_3 - t_2$, the ratio of these two times is given by

$$\frac{t_2 - t_1}{t_3 - t_2} = \frac{\epsilon (t_3 - t_1)}{(1 - \epsilon)(t_3 - t_1)} = \frac{\epsilon}{1 - \epsilon}.$$

And the corresponding ratio of the one-way outgoing and return *velocities* is therefore $\frac{1 - \epsilon}{\epsilon}$. Couldn't one argue, therefore, that the choice of a value of ϵ is not a matter of convention after all, the ground being that only the value $\epsilon = 1/2$ accords with such physical facts as the isotropy of space and with the methodological requirement of inductive simplicity, which enjoins us to account for observed facts by minimum postulational commitments? And, in any case, must we not admit (at least among friends!) that the value $1/2$ is more "true"? The contention that either the isotropy of space or Occam's "razor" is relevant here is profoundly in error, and its advocacy arises from a failure to understand the import of Einstein's statement that "we establish *by definition* that the 'time' required by light to travel from A to B equals the 'time' it requires to travel from B to A."⁵ For, in the first place, since no statement concerning a one-way transit time or one-way velocity derives its meaning from mere facts but also requires a prior *stipulation* of the criterion of clock synchronization,⁶ a choice of $\epsilon \neq 1/2$, which renders the transit times (velocities) of light in opposite directions *unequal*, cannot possibly conflict with such physical isotropies and symmetries as prevail *independently* of our descriptive conventions.⁷

⁵ A. Einstein: *PR*, *op. cit.*, p. 40.

⁶ In the case of some one-way measurements of the velocity of light such as the one by Roemer, the synchronization rule is introduced sufficiently tacitly not to be obviously present: cf. H. Reichenbach's account of Roemer's determination in "Planetenuhr und Einsteinsche Gleichzeitigkeit," *Zeitschrift für Physik*, Vol. XXXIII (1925), pp. 628–31.

⁷ An example of such an independent kind of isotropy is the fact, discovered by Fizeau, that if, in a given inertial system, the emission of a light beam traversing a *closed* polygon in a given direction coincides with the emission of a beam traveling in the opposite direction, then the returns of these beams to their common point of emission will likewise coincide. We know from the experiments by Sagnac and by Michelson and Gale [cf. A. A. Michelson and H. G. Gale: *Astrophysical Journal*, Vol. LXI (1925), p. 140, and M. G. Trocheris: *Philosophical Magazine*, Vol. XL (1949), p. 1143] that in a *rotating* system this directional symmetry will *not* obtain.

And, in the second place, the canon of simplicity which we are pledged to observe in all of the inductive sciences is not implemented any better by the choice $\epsilon = 1/2$ than by any one of the other allowed fractional values; for no distinct hypothesis concerning physical facts is made by the choice $\epsilon = 1/2$ as against one of the other permissible values. There can therefore be no question of having propounded a hypothesis making lesser postulational commitments by that special choice. On the contrary, it is the postulated fact that light is the fastest signal which assures that *each one* of the permissible values of ϵ will be equally compatible with all possible matters of fact which are independent of how we decide to set the clock at P_2 . Thus, the value $\epsilon = 1/2$ is not simpler than the other values in the inductive sense of assuming less in order to account for our observational data, but only in the descriptive sense of providing a *symbolically* simpler representation of the theory explaining these data. But this greater symbolic simplicity arising from the value $1/2$ expresses itself not only in the equality of the velocities of light in opposite directions, the ratio of the outgoing and return velocities being in general $(1-\epsilon)/\epsilon$; the value $\epsilon = 1/2$ also shows its unique descriptive advantages by assuring that synchronism will be both a *symmetric* and a *transitive* relation upon using *different* clocks in the same system: the synchronism based on $\epsilon = 1/2$ is *symmetric* as between two clocks A and B, because the setting of clock B from A in accord with $\epsilon = 1/2$ issues automatically in the fact that clock A will have the readings required by a setting from clock B in accord with $\epsilon = 1/2$, and, *mutatis mutandis* for the transitivity of the synchronism of three clocks A, B, and C.⁸ For values of $\epsilon \neq 1/2$, the symmetry and transitivity of synchronism generally do *not* obtain.

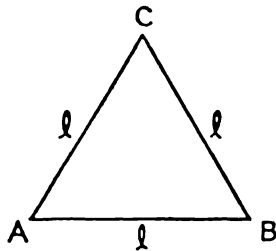
We have already refuted the ordinary language objection against using the term "topologically simultaneous" to name a relation which is not transitive. It is obvious that precisely the same refutation applies to such a linguistic objection against the philosophical admissibility of synchronization rules based on $\epsilon \neq 1/2$, which issue in a synchronism that is neither symmetric nor transitive.

I shall now rebut a final objection to the philosophical ad-

⁸ Cf. H. Reichenbach: *Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, pp. 34-35 and 38-39.

missibility of values $\epsilon \neq 1/2$ before showing the following by reference to Einstein's lightning bolt experiment: if we avail ourselves of the latitude for synchronization rules given us by the philosophical admissibility of values $\epsilon \neq 1/2$, then the physical facts postulated by the STR do not dictate discordant judgments by different Galilean frames concerning the simultaneity of given events.

To state the objection whose untenability I wish to demonstrate, I first call attention to the following results. It is a presumed empirical fact of optics in the STR that if in an inertial system, the *round-trip* time of light for an open path of one-way length l is $2T$, then the round-trip time for a *closed* path of total length nl is $n/2$ times $2T$ independently of direction along the closed path. Thus, for example, in the equilateral triangle ABC whose sides are each of length l , the time increment on the clock at A will be $3T$ for a counter-clockwise or a clockwise closed path journey ABCA (or ACBA) of a light ray which begins and ends at A.



Assume that a light ray departs from A when the clock at A reads $t = 0$. Then that clock will of course read $t = 3T$ upon the return of that light ray to A quite independently of the way in which we have synchronized (set) the clocks at B and at C. Now suppose that we use $\epsilon = 1/2$ to synchronize the B clock with the one at A, and also to synchronize the C clock with the B clock. In that case, a light ray leaving A at $t = 0$ and traveling in the counter-clockwise direction ABCA will reach the B clock when the latter reads the time

$$t_2 = t_1 + \epsilon (t_3 - t_1) = 0 + \frac{1}{2} (2T) = T.$$

And, by the same token, this same ray will reach C upon a reading of $2T$ on the C clock. We have noted that the A clock reads $3T$ upon the arrival there of our counter-clockwise light ray.

Accordingly we can see that *if* we wish to use $\epsilon = 1/2$ again to synchronize the A clock with the C clock, then the A clock is automatically already synchronized: a light-ray departing from the C clock when it reads $2T$ requires a one-way transit time of *one half* of $2T$, or T , to reach the point A, if $\epsilon = 1/2$ is used to synchronize the A clock with the one at C.

But now suppose that instead of $\epsilon = 1/2$, we had used $\epsilon = 2/3$ to synchronize the B clock with the one at A, and the C clock with the one at B. In that case, the arrival time of the counter-clockwise light ray at B would have been

$$t_2 = t_1 + \epsilon (t_3 - t_1) = 0 + \frac{2}{3} (2T) = \frac{4}{3} T.$$

And, by the same token, the arrival time of the ray at C would have been $\frac{2}{3} T$. But since the previous setting of the A clock issues in a reading of $3T$ or $\frac{5}{3} T$ on it upon the return of the light ray to A, it is clear that the one-way transit time for the ray's journey from C to A was only $\frac{1}{3} T$. And this means that the value of ϵ which was used to make the A clock synchronous with the one at C is given by

$$\frac{1}{3} T = \epsilon(2T)$$

or

$$\epsilon = 1/6.$$

Clearly, in that case the one-way velocity of light in the direction from C to A is four times greater than along either AB or BC.

In view of these results, it has been argued that the latter "compensatory" increase in the one-way velocity of light along CA is inductively highly improbable, thereby allegedly showing that on *factual* grounds of truth and falsity the value $\epsilon = 2/3$ is impermissible for setting the clocks at B and at C while the value $\epsilon = 1/2$ is uniquely warranted factually. But this argument is totally unsound, since the concept of inductive improbability of a compensation or adjustment in the one-way velocity of light along CA has no relevance here at all. For the argument is predicated on the *false* supposition that there is a factually true simultaneity of events at C and A such that there is a corresponding factually "true" synchronization of clocks at these points and a corresponding factual one-way velocity behavior of light along CA. But the one-way velocity of light depends on the criterion of simultaneity, which is conventional. And hence it is a sheer

petitio principii to offer a *reductio ad absurdum* of the conventionality of simultaneity on the alleged ground that a particular one-way velocity of light along CA is inductively improbable. The “compensation” in the one-way velocity of light along the final leg of the closed path journey of the light ray can therefore be no more inductively improbable than the following: the hundredfold increase in the numbers representing the extensions of all bodies when we shift from meters to centimeters as units of length! It can be no more false or inductively improbable that the one-way velocities of light along AB, BC, and CA correspond to the values $\epsilon = 2/3$, $\epsilon = 2/3$, and $\epsilon = 1/6$ respectively rather than to $\epsilon = 1/2$ than it is that the lengths of bodies are those corresponding to centimeters rather than to meters! Just as the convention of using centimeters after having used meters guarantees the hundredfold increase of all length numbers independently of any inductive probabilities (or improbabilities), so also the dependence of the one-way velocity along CA on the ϵ -setting of the clock at A *automatically assures* precisely the one-way velocity appropriate to that setting!

Since clock synchronizations employing values $\epsilon \neq 1/2$ are just as admissible philosophically as those employing $\epsilon = 1/2$, I shall now show that the physical facts postulated by the STR do not dictate *discordant* judgments by different Galilean frames concerning the simultaneity of given events. To furnish the required proof, consider Einstein’s familiar lightning bolt experiment in which two vertical bolts of lightning strike a moving train at points A’ and B’ respectively, and the ground at points A and B respectively, A and A’ being to the left of B and B’, respectively. Let the distance AB in the ground system be $2d$. If we now define simultaneity in the ground system by the choice $\epsilon = 1/2$, then the velocity of light traveling from A to the midpoint C of AB becomes equal to the velocity of light traveling in the opposite direction BC.⁹ And if the arrivals of these two oppositely directed light pulses coincide at C, then the ground observer will say that the bolt at AA’ struck simultaneously with the bolt

⁹ The choice of $\epsilon = 1/2$ is *not* the only choice which would assure the equality of the velocities in the opposite directions AC and BC. The same result could be achieved by using any other *equal* values of ϵ ($0 < \epsilon < 1$) to synchronize the clocks at A and B with the clock at C, although, in that case, the velocity along CA would *not* equal the velocity along AC, and similarly for the velocities along CB and BC.

at BB' , say at $t = 0$. Will the train observer's observations of the lightning flashes *compel* him to say that the bolts did *not* strike simultaneously? Decidedly not! To be sure, since the train is moving, say, to the left relatively to the ground, it will be impossible that the two horizontal pulses which meet at the midpoint C in the ground system also meet at the midpoint C' of the train system. For C' was adjacent to the point C when the ground-system clock at C read $t = 0$ and the train clock at C' read $t' = 0$, C and C' being the respective origins of the spatial coordinates.¹ Hence the flash from A' will arrive at C' *earlier* than the one from B' . But this fact will hardly require the train observer to say that the point A' was struck by lightning earlier than the point B' , *unless* the train observer *too* has chosen *equal* values of ϵ such as $1/2$ to set the clocks at A' and at B' to be each synchronized with the clock at C' in his own system, a choice which then commits him to say that the time required by light from the left to traverse the distance $A'C'$ is *the same* as the transit time of light from the right for an equal distance $B'C'$.² But, *apart from descriptive convenience*, what is to prevent the train observer from choosing suitably *unequal* values of ϵ ($0 < \epsilon < 1$) such that he too will say that the lightning bolts struck simultaneously, just as the ground observer did?

As I shall now show, the answer to this question is: nothing at all. To justify my contention in concrete detail, I shall demonstrate the following:

(i) the time-difference $\Delta t'_{C'}$ on the local clock at C' between the arrivals there of the light pulses from A' and B' is

$$\Delta t'_{C'} = \frac{2d \cdot \beta}{c \sqrt{1 - \beta^2}}.$$

(ii) If in the train system we choose values $\epsilon_{A'}$ and $\epsilon_{B'}$ ($0 < \epsilon < 1$) that satisfy the equation

$$\epsilon_{A'} - \epsilon_{B'} = \beta \quad \left(\text{where } \beta \equiv \frac{v}{c}\right)$$

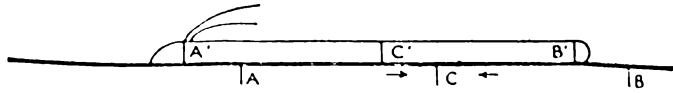
¹ A statement of the reason for postulating this adjacence of C and C' can be found in P. Bergmann: *Introduction to the Theory of Relativity* (New York: Prentice Hall; 1946), p. 31, n. 1 and p. 33.

² If a train observer moving to the *right* relatively to the ground chooses the same ϵ for synchronizing the clocks at A' and at B' with the one at C' , he will, of course, say that point A' was struck *later* than B' .

to synchronize the clocks at A' and at B' respectively with the clock at C', then the lightning flashes at A' and B' will be simultaneous in the train system while *also* being simultaneous in the ground system.

(i) The time difference $\Delta t'_{C'}$ on the *one* clock at C' is an objective fact independent of any choices of ϵ which we may make to synchronize the A' and B' clocks with the one at C'. Now the Lorentz transformation equations of the STR are based on the choice of $\epsilon = 1/2$ for all synchronizations in any given frame. But in view of the independence of $\Delta t'_{C'}$ from ϵ , we may use these Lorentz equations to compute the value of $\Delta t'_{C'}$ *without* any prejudice to our own subsequent choice of ϵ values other than 1/2 to synchronize a particular train system clock with the one at C' in part (ii) of our present proof.

Accordingly, looking at the diagram,



we have the following results. When the pulse from A' reaches C' from the left, the ground system clock at a point P adjacent to C' reads a time t given by the distance equation

$$ct + vt = d$$

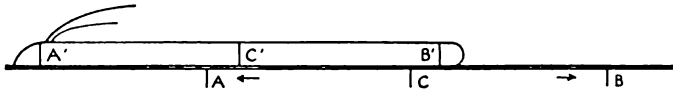
or

$$t = \frac{d}{c + v}.$$

But, in virtue of the relativistic clock retardation given by the Lorentz transformations, the train clock at C' is slow by a factor of $\sqrt{1 - \beta^2}$ as compared to the ground clock at P. Hence, upon the arrival of the A' flash at C', the clock at C' reads the time

$$t'_{C'} = \frac{d}{c + v} \sqrt{1 - \beta^2}.$$

To find the train time on the C' clock when the flash from B' reaches C', we first calculate the time reading for that event on a ground system clock at the point Q which is then adjacent to C'.



This time t is given by the following simultaneous equations, which express the fact that a light ray departing from the point B at $t = 0$ with velocity c on the ground will catch up at a distance x from B with the point C' , which departed from C at $t = 0$ in the same direction:

$$x = ct$$

and

$$x = d + vt,$$

so that

$$ct = d + vt,$$

or

$$t = \frac{d}{c - v}.$$

But again, the C' clock reading will be slow with respect to this reading on the ground system clock at Q. Hence, upon the arrival of the B' flash at C' , the clock at C' reads

$$t'_{C'} = \frac{d}{c - v} \sqrt{1 - \beta^2}.$$

Accordingly, the difference on the train clock at C' between the earlier arrival of the flash from A' and the later arrival of the one from B' is given by

$$\Delta t'_{C'} = \frac{d}{c - v} \sqrt{1 - \beta^2} - \frac{d}{c + v} \sqrt{1 - \beta^2} = \frac{2d \cdot \beta}{c \sqrt{1 - \beta^2}},$$

where $\beta \equiv v/c$, which is less than 1.

(ii) We must now ask: what are the respective values of $\epsilon_{A'}$ and $\epsilon_{B'}$ which must be used to synchronize the respective clocks at A' and B' from C' such that in the train system the two bolts struck *simultaneously* at A' and B' ? Clearly, the required values of ϵ are those which will yield the required difference $\Delta t'_{C'}$ in the arrivals at C' by yielding suitably *unequal* one-way transit times for the respective flash transits from A' to C' and from B' to C' .

To compute the required values $\epsilon_{A'}$ and $\epsilon_{B'}$, we must first recall that if T is the round-trip time of a light ray for the equal closed paths $C'A'C'$ or $C'B'C'$, then the one-way transit times for a light ray along $B'C'$ and $A'C'$ are respectively

$$(1 - \epsilon_{B'})T$$

and

$$(1 - \epsilon_{A'})T.$$

Now let τ be the time on the A' clock and *also* the time on the B' clock at which the two bolts struck *simultaneously* in the train system. Then the one-way transit time for the flash along $B'C'$ is also given by the difference between the departure time

τ and the arrival time $\frac{d}{c - v} \sqrt{1 - \beta^2}$ at C' , while the corresponding one-way transit time for the light ray journey along $A'C'$ is given by the difference between τ and the arrival time $\frac{d}{c + v} \sqrt{1 - \beta^2}$ at C' . We therefore have

$$\frac{d}{c - v} \sqrt{1 - \beta^2} - \tau = (1 - \epsilon_{B'})T$$

for the light ray journey along $B'C'$, and

$$\frac{d}{c + v} \sqrt{1 - \beta^2} - \tau = (1 - \epsilon_{A'})T$$

for the light ray journey along $A'C'$. Subtraction of the second of these equations from the first eliminates the quantity τ on the left-hand side, leaving only the difference $\Delta t'_{C'}$ on that side, so that we have

$$\frac{2d \cdot \beta}{c \sqrt{1 - \beta^2}} = (\epsilon_{A'} - \epsilon_{B'})T.$$

But since the equal round-trip distances $C'A'C'$ and $C'B'C'$ are each $\frac{2d}{\sqrt{1 - \beta^2}}$, as we know from the Lorentz transformations, the corresponding round-trip time T must be $T = \frac{2d}{c \sqrt{1 - \beta^2}}$. Substituting this result in our preceding equation, we obtain

$$\beta = \epsilon_{A'} - \epsilon_{B'} \quad Q.E.D.$$

And using any positive fractional values of ϵ compatible with the latter equation as well as the stated value of T , either of our earlier equations containing the time τ of the *simultaneous* occurrence of the bolts will permit us to solve for the value of τ

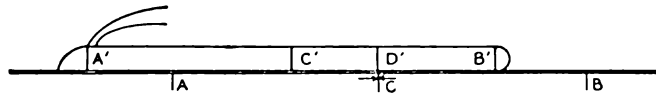
appropriate to our particular choices of ϵ . Thus, if $\beta < 1/2$ and we choose $\epsilon_{B'} = 1/2$, then $\epsilon_{A'}$ will have the fractional value $\frac{1}{2} + \beta$, and we obtain

$$\tau = \frac{d \cdot v}{c^2 \sqrt{1 - \beta^2}}$$

for the time of the *simultaneous* origination of the lightning bolt flashes at A' and B' in the train system. We see that, on the indicated *unequal* choices of $\epsilon_{A'}$ and $\epsilon_{B'}$, the train observer at C' can judge the bolts at A' and B' to have struck simultaneously,

because he would account for the time difference $\frac{2d \cdot \beta}{c \sqrt{1 - \beta^2}}$ between the arrivals of the light pulses from them at C' by the *corresponding difference* between the *one-way velocities* of the pulses along $A'C'$ and $B'C'$ respectively.

It is worthy of notice that this judgment of simultaneity in the train system need not be confined to a train observer at the midpoint C' of $A'B'$. To see that this is so, we point out first that since by hypothesis, the light pulses originating at the lightning flashes at A' and at B' will meet at the midpoint C in the ground system, they will also meet at a point D' on the train adjacent to C and lying between C' and B' , as shown on the diagram.



Let us now compute the distance of the point of encounter D' from both B' and A' from the relevant Lorentz transformation equation

$$x' = \frac{x + vt}{\sqrt{1 - \beta^2}}$$

Now the x -coordinate of D' is that of the origin at C , and the ground system time t of encounter of the two flashes at C and D' is given by $t = d/c$. Hence, inserting this value of t and $x = 0$ in the Lorentz equation, we obtain for the distance $\overline{C'D'}$

$$\overline{C'D'} = x' = \frac{vd}{c \sqrt{1 - \beta^2}}$$

By the Lorentz transformations, the distance $\overline{C'B'}$ on the train is

$$\overline{C'B'} = \frac{d}{\sqrt{1 - \beta^2}}.$$

Accordingly, subtracting $\overline{C'D'}$ from $\overline{C'B'}$ and remembering that $\beta \equiv v/c$, we obtain for the distance $\overline{B'D'}$

$$\overline{B'D'} = \frac{d}{\sqrt{1 - \beta^2}} - \frac{v}{c} \frac{d}{\sqrt{1 - \beta^2}} = d \sqrt{\frac{c - v}{c + v}}.$$

Now the distance $\overline{A'B'}$ is simply twice that of $\overline{C'B'}$, and upon subtracting from $\overline{A'B'}$ the distance $\overline{B'D'}$, we find that the distance $\overline{A'D'}$ is

$$\overline{A'D'} = \frac{2d}{\sqrt{1 - \beta^2}} - d \sqrt{\frac{c - v}{c + v}} = d \sqrt{\frac{c + v}{c - v}}.$$

Accordingly, the *ratio* of the distance $\overline{A'D'}$ to $\overline{B'D'}$ is

$$(c + v)/(c - v).$$

To assure the *simultaneity* of the flashes at A' and B' from the standpoint of a train observer at the non-midpoint D' , we need only stipulate that the ratio of the one-way velocity of light along $A'D'$ to the one-way velocity of light along $B'D'$ is the *same* as the ratio $(c + v)/(c - v)$ of the distances $\overline{A'D'}$ and $\overline{B'D'}$, thereby guaranteeing that the flashes required equal transit times to traverse these unequal paths. And the required one-way velocities along $A'D'$ and $B'D'$ can be assured by merely making appropriate choices of $\epsilon_{A'}$ and $\epsilon_{B'}$ for synchronizing the clocks at A' and at B' with the clock at D' . This can be done in an infinitude of different ways to each of which corresponds, for a given setting of the clock at D' , a particular time of the *simultaneous* occurrence of the two bolts as judged by the train observer at D' . Thus, if we again choose $\epsilon_{B'} = 1/2$ (as we did when synchronizing the B' clock with the one at C'), then the one-way velocity of light along $B'D'$ is c and hence the one-way transit time along $B'D'$ is

$$\Delta t'_{B'D'} = \frac{d}{c} \sqrt{\frac{c - v}{c + v}}.$$

Since the round-trip time of light for the closed path $D'A'D'$ is $\frac{2d}{c} \sqrt{\frac{c+v}{c-v}}$, the one-way transit time $\Delta t'_{A'D'}$ along $A'D'$ is given by

$$\Delta t'_{A'D'} = (1 - \epsilon_{A'}) \frac{2d}{c} \sqrt{\frac{c+v}{c-v}}.$$

Since we wish to choose a value of $\epsilon_{A'}$ which will result in $\Delta t'_{A'D'} = \Delta t'_{B'D'}$, we have:

$$(1 - \epsilon_{A'}) \frac{2d}{c} \sqrt{\frac{c+v}{c-v}} = \frac{d}{c} \sqrt{\frac{c-v}{c+v}}.$$

Solving for $\epsilon_{A'}$, we find

$$\epsilon_{A'} = \frac{c + 3v}{2(c + v)}.$$

To see that this value of $\epsilon_{A'}$ meets the requirement of being between 0 and 1 in the relativistic context of $v < c$, we simply rewrite $\epsilon_{A'}$ in the form

$$\epsilon_{A'} = \frac{c + 2v + v}{c + 2v + c}.$$

If the clock at D' is set in accord with the Lorentz transformations to read a time $t' = \frac{d}{c \sqrt{1 - \beta^2}}$ upon the joint arrival of the flashes from A' and B' , then the stated choices of $\epsilon_{A'}$ and $\epsilon_{B'}$ by the train observer at D' will have the following result: just as the train observer at C' , so also the train observer at D' will synchronize the A' and B' clocks with his own so as to assign the time $t' = \frac{d \cdot v}{c^2 \sqrt{1 - \beta^2}}$ to *each* of the lightning strokes at A' and B' , thereby judging them to be simultaneous. So much for the possibility of having *concordant* judgments of simultaneity as between the ground system and the train system.

The flash at AA' is causally absolutely non-connectible with any of those events at point B in the ground system which are members of the open time interval I of events bisected by the flash at BB' and of duration $4d/c$ units of time in the ground system. It is this absolute causal non-connectibility that enables the ground observer to say that the BB' flash was simultaneous with the AA' flash, as required by his choice of $\epsilon = 1/2$, and

also enables the train observer, if he wishes, to choose $\epsilon = 1/2$ for himself as well and to accept the following *discordance* of simultaneity judgments as a consequence of his own separate choice of $\epsilon = 1/2$: First, it is *not* the flash at BB' that is simultaneous with the AA' flash, as the ground-system observer maintains by his assignment of $t = 0$ to both events, but a different I-interval event E_* at B which is absolutely earlier than the BB' flash, the ground-system clock at B reading the time $t = -2vd/c^2$ upon the occurrence of E_* ; and second, the train-system time of the AA' flash and of the event E_* is

$$t' = - \frac{vd}{c^2 \sqrt{1 - \beta^2}},$$

but the BB' flash occurs at the later train-system time

$$t' = + \frac{vd}{c^2 \sqrt{1 - \beta^2}}.$$

It follows that the conventionality of simultaneity within each inertial system entailed by the aforementioned absolute causal nonconnectibility allows each Galilean observer to choose his own values of ϵ *either* so as to agree with the other observers on simultaneity *or* so as to disagree. If each separate Galilean frame does choose the value $1/2$ —which, as we saw, Einstein assumed in the formulation of the theory for the sake of the resulting descriptive simplicity—then the relative motion of these frames indeed makes disagreement in their judgments of simultaneity unavoidable, except in regard to pairs of events lying in the planes perpendicular to the direction of their relative motion.³ But, in the first instance, it is not the relative motion of inertial systems but the limiting character of the velocity of light and the behavior of clocks under transport which give rise to the relativity of simultaneity. For the Newtonian theory affirmed absolute simultaneity while countenancing the relative motion of inertial systems.

Moreover the presumed physical facts which give rise to the

³ Since $t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \beta^2}}$, spatially separated events occurring at the same time t in system S will occur at different times t' in system S' , unless they lie in the planes given by $x = \text{constant}$. Cf. H. Reichenbach: *Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, pp. 55–56.

conventionality and relativity of simultaneity (as distinct from man's discovery of these facts) are evidently quite independent of man's presence in the cosmos and of his measuring activities. Hence, the relativity of simultaneity is not first generated by our inability to carry out those measuring operations that would yield relations of absolute simultaneity in the Newtonian sense. Instead, the failure of human signaling and measuring operations to disclose relations of absolute simultaneity is only the epistemic consequence of the primary nonexistence of these relations. It is therefore erroneous or at best highly misleading to give an epistemic or homocentrically operational twist to Einstein's conception of simultaneity by emphasizing the role of signals as a means of human knowledge. Such a homocentrically operational twist falsely portrays what Einstein construes, in the first instance, as an ontological feature of the temporal relatedness of physical events. For the homocentrically operational account suggests that Einstein's repudiation of Newton's absolute simultaneity rests on a mere *epistemic* limitation on the *ascertainability* of the existence of relations of absolute simultaneity. To be sure, human operations of measurement are indispensable for *discovering* or *knowing* the physical relations and thereby the time relations sustained by particular events. But these relations are or are not sustained by physical macro-events quite apart from *our* actual or hypothetical measuring operations and are *not* first *conferred* on nature by our operations. In short, it is because no relations of absolute simultaneity *exist* to be measured that measurement cannot disclose them; it is not the mere failure of measurement to disclose them that *constitutes* their nonexistence, much as that failure is *evidence* for their nonexistence.

Only a philosophical obfuscation of this state of affairs can make plausible the view that the relativity of simultaneity (or, for that matter, any of the other philosophical innovations of relativity theory) lends support to the subjectivism of homocentric operationism or of phenomenalistic positivism.⁴

⁴ For further details on why I deny that the STR supports an operational account of scientific concepts in P. W. Bridgman's homocentric sense, see A. Grünbaum: "Operationism and Relativity," *The Scientific Monthly*, Vol. LXXIX (1954), p. 228; reprinted in *The Validation of Scientific Theories*, *op. cit.*, pp. 84-94.

(C) HISTORY OF EINSTEIN'S ENUNCIATION OF THE LIMITING CHARACTER OF THE VELOCITY OF LIGHT IN VACUO.

We can restate what we have called assumption (ii) above as follows:

(ii) light is the fastest signal *in vacuo* in the following topological sense: no kind of causal chain (moving particles, radiation) emitted *in vacuo* at a given point A together with a light pulse can reach any other point B earlier, as judged by a local clock at B which merely orders events there in a metrically arbitrary fashion, than this light pulse.⁵

We saw that this assumption (ii)—along with Einstein's postulate that clock transport does *not* define relations of absolute simultaneity, which we called assumption (i)—is fundamental to Einstein's conception of simultaneity. It therefore behooves us to inquire in the present section into the grounds which Einstein had in 1905 for making assumption (ii). And we note, incidentally, that our very awareness of the fundamental historical problem we have just posed depends upon a clear prior philosophical comprehension of Einstein's conception of simultaneity, free from the prevalent misunderstandings which we have criticized.

As will presently be evident, Einstein himself gives us tantalizingly incomplete explicit information concerning the grounds for his original confidence in his intuition that assumption (ii)

⁵ It must be remembered that, in its metrical form, this postulate does not confine purely "geometric" velocities which are not the velocities of causal chains to the value c . For example, within any given inertial system K , special relativity allows us to subtract vectorially (as against by means of the Einstein velocity addition formula) a velocity $v_1 = .9c$ of body A in K and a velocity $v_2 = -.3c$ of body B in K to obtain $1.2c$ as the relative velocity of separation of A and B as judged by the K -observer [though not by an observer attached to either A or B]. For no body or disturbance is thereby asserted to be traveling relative to any other at the velocity $1.2c$ of separation, and since the process of separation has no direction of propagation, it cannot be a causal chain. See Max von Laue: *Die Relativitätstheorie* (Braunschweig: F. Vieweg; 1952), Vol. I, pp. 40–41; L. Silberstein: *The Theory of Relativity* (New York: The Macmillan Company; 1914), p. 164n; A. Sommerfeld: Notes in *PR*, a collection of original memoirs (New York: Dover Publications, Inc.; reprint), p. 94; A. Grünbaum: *Philosophy of Science*, Vol. XXII (1955), p. 53; H. P. Robertson: *Mathematical Reviews*, Vol. XVI (1955), p. 1166; G. D. Birkhoff: *The Origin, Nature and Influence of Relativity* (New York: The Macmillan Company; 1925), p. 104, and J. Weber: *American Journal of Physics*, Vol. XXII (1954), p. 618.

is true. But even if we did possess full clarity on that score, we still confront the same question in regard to assumption (i) and are compelled to try to answer it with even less assistance from Einstein himself, as we shall see.

Correction in
Append. §33

The importance of also understanding the grounds on which Einstein thought he could safely make assumption (i) can be gauged by the following basic fact: if, in accord with Newtonian conceptions, assumption (i) had been thought to be *false*, then the belief in the truth of (ii) alone would *not* have warranted the abandonment of the received Newtonian doctrine of absolute simultaneity. And, in that eventuality, the members of the scientific community to whom Einstein addressed his paper of 1905 would have been fully entitled to *reject* his *conventionalist* conception of *one-way* transit times and velocities. But the denial of absolute simultaneity by the latter conception is *crucial* for his principle of the constancy of the speed of light, as is evident from his statement of this principle in Section 2 of his 1905 paper. Specifically, Einstein emphasizes there that if the *one-way* velocity of light is also to have the numerical value c , then "time interval is to be taken in the sense of the [convention or] definition [of simultaneity] in Section 1."⁶ The indispensability of assumption (i) for Einstein's conception of simultaneity is further apparent from the fact that a Newtonian physicist quite naturally regards *not* signal connectibility but the readings of suitably transported clocks as the fundamental indicators of temporal order. The Newtonian recognizes, of course, that the truth of (ii) compels such far-reaching revisions in his theoretical edifice as the repudiation of his second law of motion, which allows particles to attain arbitrarily high velocities through accelerations of appropriate durations. But he stoutly and rightly maintains that *if* (i) is *false*, absolute simultaneity remains intact, unencumbered by the truth of (ii).

It is of both historical and logical importance to note that in "The Principles of Mathematical Physics," an address delivered in St. Louis on September 24, 1904, and published in *The Monist*, Vol. XV (1905), pp. 1-24,⁷ H. Poincaré had already envisioned

⁶ Cf. A. Einstein: *PR*, *op. cit.*, p. 41.

⁷ This address by Poincaré is reprinted in the *Scientific Monthly*, Vol. LXXXII (1956), p. 165.

the construction of a new mechanics in which the velocity of light would play a limiting role. Although he concluded this paper with the remark that "as yet nothing proves that the (old) principles will not come forth from the combat victorious and intact,"⁸ Poincaré did prophesy that:

From all these results, if they are confirmed, would arise an entirely new mechanics, which would be, above all, characterized by this fact, that no velocity could surpass that of light, any more than any temperature, could fall below the zero absolute, because bodies would oppose an increasing inertia to the causes, which would tend to accelerate their motion; and this inertia would become infinite when one approached the velocity of light.⁹

But, unlike Einstein, Poincaré failed to see the import of this new limiting postulate for the meaning of simultaneity: in the very same paper,¹ he speaks of spatially separated clocks as marking *the same hour at the same physical instant* (my italics). And he distinguishes between watches which mark "the true time" and those that mark only the "local time," a distinction which was also invoked, as we shall see, by Larmor and Lorentz. But Poincaré does note that a Galilean observer will not be able to detect whether his frame is one in which the clocks mark the allegedly spurious local time. Accordingly, he affirms "The principle of relativity, according to which the laws of physical phenomena should be the same, whether for an observer fixed, or for an observer carried along in a uniform movement of translation; so that we have not any means of discerning whether or not we are carried along in such a motion."²

Specifically what does Einstein himself tell us about his original grounds for assuming the truth of assumption (ii)? At this point, it is essential to quote him *in extenso*. He writes:

By and by I despaired of the possibility of discovering the true laws by means of constructive efforts based on known

⁸ H. Poincaré: "The Principles of Mathematical Physics," *The Monist*, Vol. XV (1905), p. 24.

⁹ *Ibid.*, p. 16.

¹ *Ibid.*, p. 11.

² *Ibid.*, p. 5.

For a useful, brief account of these pre-Einsteinian conceptions, cf. P. Bergmann: "Fifty Years of Relativity," *Science*, Vol. CXXIII (1956), p. 123.

facts. The longer and the more despairingly I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. The example I saw before me was thermodynamics. The general principle was there given in the theorem: the laws of nature are such that it is impossible to construct a *perpetuum mobile* (of the first and second kind). How, then could such a universal principle be found? After ten years of reflection such a principle resulted from a paradox upon which I had already hit at the age of sixteen: If I pursue a beam of light with a velocity c (velocity of light in a vacuum), I should observe such a beam of light as a spatially oscillatory electromagnetic field at rest. However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations. From the very beginning it appeared to me intuitively clear that, judged from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer who, relative to the earth, was at rest. For how, otherwise, should the first observer know, i.e., be able to determine, that he is in a state of fast uniform motion?

One sees that in this paradox the germ of the special relativity theory is already contained. Today everyone knows, of course, that all attempts to clarify this paradox satisfactorily were condemned to failure as long as the axiom of the absolute character of time, viz., of simultaneity, unrecognizedly was anchored in the unconscious. Clearly to recognize this axiom and its arbitrary character really implies already the solution of the problem. The type of critical reasoning which was required for the discovery of this central point was decisively furthered, in my case, especially by the reading of David Hume's and Ernst Mach's philosophical writings.³

We see that Einstein gives essentially three reasons for his original belief in assumption (ii), noting that, like two of the laws of thermodynamics, this assumption is a "principle of impotence," to use E. T. Whittaker's locution.⁴ Einstein's seemingly distinct three reasons are the following: (1) "on the basis of experience," there are no "stationary" light waves, (2) Neither

³ A. Einstein: "Autobiographical Notes," *op. cit.*, p. 53.

⁴ Cf. E. T. Whittaker: *From Euclid to Eddington* (Cambridge: Cambridge University Press; 1949), §25, pp. 58-60.

are there any such phenomena on the basis of Maxwell's equations, and (3) At the very beginning, there was intuitive clarity that preferred inertial systems do not exist, the laws of physics, including those of light propagation, being the same in all of them. These three reasons invite the following corresponding three comments.

1. The failure of our experience to have disclosed the existence of stationary light waves is not, of course, presumptive of their non-existence, unless that experience included the circumstances requisite to our observation of such waves if they do exist. What could such circumstances be? Suppose it were physically possible for a star to recede from the earth at the speed c of light. Assuming that there actually is such a star, postulate further that the speed of the light emitted by the star in the direction of the earth is c relatively to the *star*, the light's speed relatively to the earth being given by the *Galilean-Newtonian* velocity addition and hence being *zero*. Then the earth would maintain a *constant distance* from the light wave. And if there were a way for us to register the presence of that stationary light wave, then we could have evidence of its existence. *Mutatis mutandis*, a light source in the laboratory moving at the velocity c might have produced the same kind of phenomenon.

Perhaps Einstein envisaged these kinds of conditions as situations in which our experience ought to have disclosed the existence of stationary light waves.

If so, one wonders, however, how much weight he actually attached to this observational argument on behalf of assumption (ii). For he was undoubtedly cognizant of the contingency of the conditions governing the observable occurrence of the phenomenon in question. In particular, it should be noted that Einstein's mention of "the basis of experience" in this context *cannot* be assumed to be referring to the 1902–1906 experiments by Kaufmann and others on the deflection of electrons (β -rays) in electric and magnetic fields. For if we suppose him to have been familiar with these experiments, they must have left him in a quandary precisely in regard to the truth of assumption (ii): while yielding a mass variation with velocity incompatible with Newtonian dynamics, the results of these experiments were un-

able to rule out the formulae of *Abraham's* dynamics, which *allowed* particle velocities *exceeding* the velocity of light in *vacuo*.⁵

John Stachel has made the interesting suggestion that the clue to Einstein's reference to "the basis of experience" may be found in his commentary on the significance of the results of Fizeau's experiments on the velocity of light in moving liquids. Writing concerning the empirical formula codifying Fizeau's results, Einstein says: "From the stated formula, one can make the interesting deduction that a liquid which does not refract light at all ($n = 1$) would not affect the propagation of light with respect to it even if the liquid is moving."⁶

2. Can stationary light waves be regarded as ruled out by Maxwell's equations, if one does not already accept the principle of relativity, which guarantees the validity of the usual form of Maxwell's equations in all inertial systems? In other words, can the *second* of Einstein's avowed reasons for his dismissal of the possibility of stationary light waves be regarded as other than a logical consequence of the third? Clearly, if Maxwell's equations *are* coupled with the principle of relativity, then these equations indeed rule out stationary light waves in every inertial system. But since Maxwell's equations are not covariant under *Galilean* transformations, it is far from clear that stationary light waves are precluded by the form assumed by the equations in an inertial system *S* moving with the velocity *c* relatively to the primary (aether) frame *K* *and* having coordinates which are related by the *Galilean* transformations to those of the *K*-system. Since Einstein does not mention the "Galilean transform" of Maxwell's equations, it would seem that the only reason why he felt justified in regarding Maxwell's equations as support for his repudiation of stationary light waves was that he had already assumed the principle of relativity on intuitive grounds.

3. In view of the presumably flimsy character of the appeal to

⁵ For details, cf. M. von Laue: *Die Relativitätstheorie*, *op. cit.*, Vol. I, pp. 26–27. Experiments which can reasonably be interpreted as strongly supporting Einstein's assumption (ii) were not successfully carried out until a good many years after 1905: Cf. C. Møller: *The Theory of Relativity*, *op. cit.*, pp. 85–89.

⁶ A. Einstein: "Die Relativitätstheorie," in: E. Warburg (ed.) *Physik* (Leipzig: Teubner; 1915), p. 704.

experience and of the redundancy of (2) with (3) among the reasons given by Einstein, we are pretty much left with his intuitive confidence in the principle of relativity as the basis for his assumption of (ii).

It is not uncommon to hear it said that Einstein first arrived at assumption (ii) by deducing it from his formulas for the addition of velocities, these formulas thereby allegedly constituting the grounds on which he first convinced himself that assumption (ii) was true. This claim is *not* supported by relevant historical evidence, but has been put forward as the *historical corollary* of an inveterate compound error in which both a formal logical blunder and a philosophical misconception are ingredient: the doubly false supposition that (a) the relativistic addition law for velocities entails assumption (ii), and that (b) assumption (ii) is *not* presupposed by the principle of the constancy of the speed of light (and hence by the Lorentz transformations) but is first derived in the STR from the law for velocity addition.

I shall now demonstrate, however, that both (a) and (b) are grossly in error.

If an object has components of velocity u_x' , u_y' and u_z' in an inertial system K' moving with the velocity v along the positive x -axis of another inertial system K , then the corresponding components of velocity u_x , u_y , and u_z in K are:

$$u_x = \frac{u_x' + v}{1 + u_x' v/c^2}$$

$$u_y = \frac{u_y' (1 - v^2/c^2)^{\frac{1}{2}}}{1 + u_x' v/c^2}$$

$$u_z = \frac{u_z' (1 - v^2/c^2)^{\frac{1}{2}}}{1 + u_x' v/c^2}$$

Whatever may be the components of velocity in one system at a given time, these transformation equations relate them to the corresponding components in the other system. This is shown by the behavior of the function $u_x = f(u_x', v)$ for values of the *independent* variable u_x' which *exceed* c , values whose exclusion from the range of u_x' clearly requires grounds *other than* the above velocity transformations. Thus, if $u_x' = 2c$ while $v = c/2$, then $u_x = 5c/4$. Again, if $v < c$, a finite velocity *greater* than c

of magnitude $u_x = c^2/v$ corresponds to an *infinite* velocity u_x' . And the super- c values of the component velocities that are allowed by the velocity-addition laws are, in fact, likewise allowed by the STR as a whole *provided* these velocities pertain to propagations in which no causal influence, energy, or matter is transmitted at super- c speeds. It is *immaterial* that the above equations for the components u_y and u_z do *rule out* that the *system* K' have a relative velocity v in excess of c by yielding imaginary values for that case. For this result does *not* preclude that accelerating objects have time-dependent *super- c* component velocities in K' .

What the relativistic addition formulas do show, and all that Einstein claimed for them in this connection in Section 5 of his 1905 paper, is that "from a composition of two velocities which are less than c , there always results a velocity less than c " and that "the velocity of light c cannot be altered by composition with a velocity less than that of light."⁷ The fallacy of inferring assumption (ii) from the relativistic velocity-addition laws dies hard, however, since it has been explicitly committed in the writings of such eminent authors as E. T. Whittaker⁸ and R. C. Tolman.⁹

It is likewise an error to suppose that assumption (ii) is *not presupposed* by the principle of the constancy of the speed of light (and hence by the Lorentz transformations), the result of this error being that assumption (ii) is then wrongly believed to still require logical justification by the time the velocity addition law is derived from the Lorentz transformations. That these suppositions are mistaken can now be readily shown as follows. Einstein deduced the velocity-addition laws in Section 5 of his 1905 paper via the mediation of the Lorentz transformations from the two basic principles of his Section 2, i.e., from the principle of relativity and the principle of the constancy of the speed of light. Now, the latter principle presupposes, as he notes *pointedly* at the start of Section 2, that the *one-way* transit-time ingredient

⁷ Cf. A. Einstein: *PR*, *op. cit.*, p. 51.

⁸ E. T. Whittaker: *A History of the Theories of Aether and Electricity* (London: Thomas Nelson & Sons, Ltd.; 1953), Vol. II, p. 38.

⁹ R. C. Tolman: *Relativity, Thermodynamics and Cosmology* (Oxford: Oxford University Press; 1934), p. 26.

in the *one-way* velocity of light is based on the *definition of simultaneity* given in his Section 1, which in turn rests on an avowedly conventionalist as opposed to absolutist conception of simultaneity. Since the deduction of the velocity-addition laws thus presupposes the *denial* of absolute simultaneity, it clearly presupposes as *one* of its premises the limiting character of the velocity of light *in vacuo*, i.e., assumption (ii). Hence, even if assumption (ii) were deducible from the velocity-addition law, this law could not have been the basis on which Einstein first convinced himself of the truth of assumption (ii).

It is apparent that logical and philosophical confusion concerning the role of assumption (ii) has begotten a historical falsehood as to how Einstein first arrived at that assumption.

We are now in a position to make a conjecture as to Einstein's grounds for assumption (i). We recall that (i) asserts that within the class of events, material clocks do not define relations of absolute simultaneity under transport. As Einstein states explicitly both in our citation from his intellectual autobiography and in his formulation of the principle of the constancy of the speed of light in Section 2 of the 1905 paper, *the absolutistic conception of simultaneity* was the Gordian knot obstructing the resolution of his boyhood paradox, i.e., the reconciliation of the two basic principles of his Section 2. But, assumption (i) was *required* no less than (ii) for the denial of absolute simultaneity! Hence his confidence in (i) must be presumed to have derived from his belief in the correctness of *both* of the two principles in his Section 2. One wonders in this connection what role, if any, was played in the development of Einstein's reflections on simultaneity by Poincaré's *obscurely* conventionalist treatment of simultaneity in his paper "La Mesure du Temps."¹

Would it be safe to conclude from Einstein's autobiographical statement that actual experimental results bearing on the velocity of light *in vacuo* and on its status as a limiting velocity in fact played no role at all when he groped his way to an espousal of the principle of relativity? If so, there would be the following serious question: can the theoretical guesses of an Einstein be regarded to have been genuinely more *educated*, as opposed to

¹ H. Poincaré: "La Mesure du Temps," *op. cit.*, pp. 1-13.

just more *lucky*, than the abortive fantasies of those quixotic scientific thinkers whose names have sunk into oblivion and whose subjective confidence in having made a bona fide discovery was no less passionate than was Einstein's? This important question, which I raised in an earlier publication,² was not answered but merely evaded by M. Polanyi, who—on the strength of the wisdom of a half century of hindsight!—offers the following *petitio principii* in criticism of me: "Dr. Grünbaum . . . seems repeatedly to express the view that Einstein had no sufficient reason to adopt the fundamental assumptions of relativity when he in fact did so. But he [Grünbaum] is more puzzled by this than I am, for he does not allow for the unspecifiable clues which justifiably [sic] guided Einstein's formulation."³ But this declaration immediately prompts the further question of just how Polanyi's account illuminates the epistemological (methodological) attributes of a bona fide scientific discovery as opposed to those of a plausible and cherished but abortive speculation, while maintaining, as he does, both the *fallibility* and the *unspecifiability* of the clues on which scientific theorists rely. Flying the banner of unspecifiability of clues while heaping scorn on the demand for adequate empirical support of knowledge claims in physics does not absolve Polanyi from the necessity of answering this further question. For example, what besides his present, inherently *ex post facto* possession of the wisdom of hindsight enables Polanyi's unspecifiability thesis to discriminate in regard to methodological justifiability between the following two hypotheses at the time of their initial enunciation: first, Einstein's assumption that the velocity of light is independent of the velocity of the emitting source, which has subsequently turned out to be successful (H. Dingle's allegations to the contrary notwithstanding),⁴ and second, Ritz's contrary hypothesis, which, to the profound disappointment and chagrin of his followers, turned

² Cf. A. Grünbaum: "The Genesis of the Special Theory of Relativity," in H. Feigl & G. Maxwell (eds.) *Current Issues in the Philosophy of Science*, *op. cit.*, p. 49.

³ M. Polanyi: "Notes on Professor Grünbaum's Observations," in: *Current Issues in the Philosophy of Science*, *op. cit.*, pp. 54–55; hereafter this paper by Polanyi will be cited under the abbreviation "Notes, etc."

⁴ Cf. A. Grünbaum: "Professor Dingle on Falsifiability: A Second Rejoinder," *British Journal for the Philosophy of Science*, Vol. XII (1961), pp. 153–56.

out to be unsuccessful after his death in 1909? If Polanyi's answer were to be that each of these contrary hypotheses was indeed methodologically justified in the context of its own set of unspecified but fallible clues, then his account altogether loses its relevance to its avowed subject. For it then throws no light whatever on the epistemology of scientific discovery as distinct from the psychology of imaginative but pathetically abortive speculation. The conclusion seems inescapable that if Polanyi is at all to avoid the well-known aprioristic pitfalls of classical rationalism, his conception of the logic of scientific discovery compels him to seek refuge in the unspecifiability of clues as an *asylum ignorantiae*. And his answer to my question as to what, besides good fortune, does distinguish *methodologically* between scientific discoveries of genius and brilliant flights of scientific fantasy merely baptizes the difficulty which I raised and gives it the name "unspecifiability."

But is it true that actual experiments on the velocity of light such as the Michelson-Morley experiment did not play any genetic role in the STR? E. T. Bell, writing *before* the publication of Einstein's "Autobiographical Notes," and referring to the influence of the Michelson-Morley experiment on Einstein, claims that "he has stated explicitly that he knew of neither the experiment nor its outcome when he had already convinced himself that the special theory was valid."⁵ And M. Polanyi reports⁶ that Einstein had authorized him in 1954 to publish the statement that "the Michelson-Morley experiment had a negligible effect on the discovery of relativity."

Yet if *both* of the reasons for the STR adduced by Einstein in the Introduction to his 1905 paper were among the factors which had prompted his initial espousal of the STR, the consonance of the foregoing claims by Bell and Polanyi with Einstein's own text there is quite problematic. For in that Introduction, Einstein cites the following two considerations as suggesting "that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest":⁷ (1)

⁵ E. T. Bell: *The Development of Mathematics* (New York: McGraw-Hill Book Company; 1945), p. 210.

⁶ M. Polanyi: *Personal Knowledge* (Chicago: University of Chicago Press; 1958), pp. 10–11, (hereafter cited under the abbreviation *PK*.)

⁷ A. Einstein: *PR*, *op. cit.*, p. 37; italics are mine.

the lack of *symmetry* in the classical electrodynamic treatment of a current-carrying wire moving relatively to a magnet at rest, on the one hand, and of a magnet moving relatively to such a wire at rest, on the other, (2) "*the unsuccessful attempts to discover any motion of the earth relatively to the 'light medium' [aether].*" Unless they provide some other consistent explanation for the presence of the latter statement in Einstein's text of 1905, it is surely incumbent upon all those historians of the STR who deny the inspirational role of the Michelson-Morley experiments to tell us specifically what other "unsuccessful attempts to discover any motion of the earth relatively to the 'light medium'" Einstein had in mind here. And it was likewise incumbent upon Holton to tell us how he reconciles this explicit reference to experiments, whose identity was presumably clear to Einstein's public in 1905, with the following claims:⁸ Einstein's paper "begins with the statement of formal asymmetries or other incongruities of a predominantly esthetic nature (rather than, for example, a puzzle posed by unexplained experimental facts)," [p. 629] and "The [1905 STR] paper does not invoke explicitly any of the several well-known experimental difficulties" [p. 630]. This obligation to take cognizance of the second consideration in the Introduction to the 1905 paper should also have been shouldered by the mature *reminiscing* Einstein himself when authorizing the statement given by Polanyi.

All the more so, since, as early as 1915, Einstein himself gave the following *historical* account in his contribution "Die Relativitätstheorie" to the volume *Physik*:⁹ "It is hardly possible to form an independent judgment of the justification of the theory of relativity, if one does not have some acquaintance with the experiences and thought processes which preceded it. Hence these will need to be discussed first. [p. 703] . . . The successes of the Lorentzian theory were so significant that the physicists would have abandoned the principle of relativity without qualms, had it not been for the availability of an important experimental result, . . . namely Michelson's experiment" [p. 706]. It is quite unfortunate that the serious *prima facie* discrepancies between Einstein's testimony of 1905 and 1915, on the one hand, and of

⁸ G. Holton: "On the Origins of the Special Theory of Relativity," *op. cit.*, p. 629 and p. 630.

⁹ A. Einstein: "Die Relativitätstheorie," *op. cit.*, pp. 703 and 706.

1954 (as transmitted to Polanyi via Dr. N. Balazs), on the other, did not restrain Polanyi from making the following assertion or at least induce him to modify it: "When Einstein discovered rationality in nature, unaided by any observation that had not been available for at least fifty years before, our positivistic textbooks promptly covered up the scandal by an appropriately embellished account of his discovery," i.e., by falsely portraying relativity "as a theoretical response to the Michelson-Morley experiment."¹

Conceivably our historical conundrum is resolved by Professor Henry S. Frank's suggestion that when writing his STR paper for the scientific public of 1905, Einstein deemed it appropriate to appeal also to the failure of the aether theory in the Michelson-Morley experiments as a justification of the STR, even though that failure had genetically not been a reason for his own initial confidence in the STR. But this suggestion leaves us just as puzzled concerning the logical, as distinct from psychological grounds which would then originally have motivated Einstein to have confidence in the principle of relativity without the partial support of the Michelson-Morley experiment, while that very lack of support would have sufficed, by his own admission, to assure the abandonment of the principle "without qualms" by his colleagues. And our puzzlement is deepened by Bernard Jaffe's report that in a letter to him, Einstein expressed his debt to the American physicist [i.e., Michelson] in these words:

It is no doubt that Michelson's experiment was of considerable influence upon my work insofar as it strengthened my conviction concerning the validity of the principle of the special theory of relativity. On the other side I was pretty much convinced of the validity of the principle before I did know this experiment and its result. In any case, Michelson's experiment removed practically any doubt about the validity of the principle in optics, and showed that a profound change of the basic concepts of physics was inevitable.²

And Jaffe adds that "In 1931, just before the death of Michelson, Einstein publicly attributed his theory to the experiment of Michelson."

¹ M. Polanyi: *PK*, *op. cit.*, p. 11.

² B. Jaffe: *Michelson and the Speed of Light* (New York: Doubleday & Company; 1960), pp. 100-101.

Oddly enough, Polanyi quotes Einstein's passage about the non-detection of the earth's aether-motion just after having declared that the outcome of the Michelson-Morley experiment was "on the basis of pure speculation, rationally intuited by Einstein before he had ever heard about it."³

On this, my comments are the following. Suppose it were clear, which it is not by any means, that Einstein's frustratingly *non-specific* mention of attempts to confirm the aether theory experimentally is to be explained by Professor H. S. Frank's suggestion and is thus consonant with Polanyi's discounting of the Michelson-Morley experiment. Then we would still have to emphasize that this historical conjecture cannot be used to vitiate an account of both the genesis and justification of scientific theory which has the following two characteristics: first, while being *empiricist*, it is broad enough to accommodate the *valid core* of the Kantian emphasis on the *active*, creative role of the scientific imagination in the postulational elaboration of hypothetico-deductive theories, and second, by being empiricist, it avoids the notorious aprioristic pitfalls of classical rationalism, pitfalls on whose brink Polanyi hovers⁴ in the protective twilight of his thesis of the unspecifiability of fallible clues. Can the history of the STR as conjectured by Polanyi be validly adduced here to prove any more than the untenability of a *crudely* empiricist Machian or Aristotelian-Thomist, abstractionist account of theory-construction as a mere codification of the results of experiments? And is the untenability of the latter kind of account not fully recognized by any empiricist conception of scientific knowledge which incorporates the lesson of Kant? Does such an empiricist conception not allow for the difference between knowing of an experimental result in a narrow sense, on the one hand, and speculatively assigning a wider significance to it, on the other?

That the answer to the last two of these questions is indeed "yes" is indicated by a single chapter title in a book by an empiricist writer whose views are presumably anathema to Polanyi: Chapter Six of Reichenbach's *Rise of Scientific Philosophy* is entitled "The Twofold Nature of Classical Physics: Its Empirical and Its Rational Aspect."

See Append. §34 for
amplification

³ Polanyi, *PK*, *op. cit.*, p. 10.

⁴ M. Polanyi: *PK*, *op. cit.*, p. 15.

In a further attempt to document the fideist nature of allegiance to physical theories, Polanyi refers to the scientific community's reception of D. C. Miller's report of *positive* effects obtained by him in his repetitions of the Michelson-Morley experiment. Thus, Polanyi cites the history of the STR to maintain further that once the "rational" appeal of the theory had captured the minds of the scientific community, contrary evidence hardly acted as a check on its subsequent acceptance. But Polanyi's indictment scores only against those who mistakenly portray evidence-reports as not already theory-laden. For it is the false absolute dichotomy between theory and evidence which would prompt an uncritically hasty *disconfirmatory* use of what is *prima facie* contrary evidence. As Duhem has emphasized, *collateral* hypotheses pertaining to the conditions of an experiment and to the laws of operation of its test equipment are an essential ingredient in the logical fabric of disconfirmation. And thus there is a mutuality of accreditation between theory and evidence in virtue of the interpenetration of the criteria of credibility which certify evidence as *bona fide*, on the one hand, and theory as evidentially warranted on the other.

It was therefore hardly a case, as Polanyi would have it, of simply ignoring negative evidence when adherents of the STR surmised—correctly as it seems to have turned out⁵—that the conditions under which Miller obtained his results were different from what he had supposed them to be. And there was also the inductively reasonable suspicion that the Cleveland site at which Miller had obtained his aberrant results was not a terrestrial singularity in regard to delicate optical phenomena, since it had not proven to be a singularity with respect to other (coarse) physical phenomena.

We must examine a final but essential point in Polanyi's critique of the account given by philosophers of science of the relevance of the discovery of the STR to an analysis of its epistemological foundations. Polanyi denies the legitimacy of the distinction between the psychology and the epistemology (logic) of discovery, a distinction which Reichenbach usefully termed as being between the "context of discovery" and the "context of

⁵ Cf. R. S. Shankland, S. W. McCuskey, S. C. Leone, and G. Kuerti: *Reviews of Modern Physics*, Vol. XXVII (1955), p. 167.

justification." Quoting H. Mehlberg's statement that "The gist of the scientific method is . . . verification and proof, not discovery," Polanyi writes: "Actually, philosophers deal extensively with induction as a method of scientific discovery; but when they occasionally realize that this is not how discoveries are made, they dispose of the facts to which their theory fails to apply by relegating them to psychology."⁶ And while denying that he intends to discount completely the role of experimental results in the genesis of acts of discovery,⁷ Polanyi does make the following claims concerning the criteria by which *bona fide* scientific discoveries can legitimately be judged as such:

Knowledge of the external world is in general acquired by relying on clues which cannot be fully identified.⁸ . . . Scientific discoveries are likewise based on clues that are never fully specifiable. This may be called "intuition" but I rarely use this term, for it is traditionally charged with the fallacious connotation of infallibility, which I do not wish to imply.⁹ . . . A good problem is a passionate intimation of a hidden truth.¹ . . . I recognize the continued operation of these anticipatory powers in the verification of discovery and holding of knowledge. The scientist's conviction of having arrived at some true knowledge is akin to the powers by which he recognises a problem. It is an anticipation of yet hidden truths, which in this case are expected to emerge in the uncertain future as the unknown consequences of the truth as known at present. When we are told that fruitfulness is the characteristic of a true discovery, the actual facts are obscured. For it is absurd to suggest that we should recognise truth by the wealth of its still unknown future consequences; on the contrary, our recognition of a true piece of knowledge is an anticipation of such unknown consequences, and if we acknowledge that knowledge is rightly held we subscribe to these anticipations. They are an expression of the belief that true knowledge is an aspect of a hidden reality which as such can yet reveal itself in an indeterminate range of future discoveries.²

⁶ M. Polanyi: *PK, op. cit.*, p. 14 n. 1.

⁷ M. Polanyi: "Notes, etc.," *op. cit.*, p. 55.

⁸ *Ibid.*, p. 53.

⁹ *Ibid.*, p. 54.

¹ *Ibid.*

² *Ibid.*

Given his affirmation of the unspecifiability of clues and his admission of their fallibility, this statement of Polanyi's succeeds no more than his earlier ones in providing a consistent articulation of the epistemological attributes of a bona fide discovery as contrasted with those of an initially plausible, passionately espoused but wholly abortive speculation. But in the absence of precisely such an articulation, what is to be our verdict on his indictment of Mehlberg, Reichenbach, and others who do invoke the distinction between the *psychology* of the propounding of scientific hypotheses, on the one hand, and the *epistemological justification* of these hypotheses on the other? It can be none other than that Polanyi's indictment is altogether gratuitous. To repeat, the anticipations and commitments of which he speaks are initially no less passionate in cases of hypotheses which turn out to be untenable than in cases of theoretical conjectures which are abundantly borne out by subsequent evidence. And hence those of Polanyi's remarks which are perceptive pertain not, as he believes, to scientific discovery; instead, they do apply after all only to the *psychology* of the propounding of scientific speculations. For what is it that makes a hypothesis *true*, and what warrants the claim that a bona fide discovery has actually been made? Surely it is not the fact that the scientist has committed himself to the as yet unknown and unverified anticipations flowing from the hypothesis! When maintaining that fruitfulness is the test of a true discovery, philosophers of science are not asserting, as Polanyi supposes incorrectly, that we recognize truth "by the wealth of its still unknown consequences." Fruitfulness in the form of both available evidence *and* future corroborations is indeed the test of the truth of a hypothesis. But we recognize the latter truth at a given time on the basis of supporting evidence then available, this recognition being understood to require the certification of credibility instead of being constituted by some ineffable, thaumaturgical insight.

We conclude against Polanyi that the philosopher's relegation of the promptings of what turn out to be lucky hunches to the psychology of creative work in science is *not* an example of the fallacy of ignoring negative instances. As well condemn the avowed neglect by the literary critic of Friedrich Schiller's re-

ported use of the smell of rotten apples as an aid to his writing of *Die Glocke*.

(D) THE PRINCIPLE OF THE CONSTANCY OF THE SPEED OF LIGHT AND THE FALSITY OF THE AETHER-THEORETIC LORENTZ-FITZGERALD CONTRACTION HYPOTHESIS.

The principle of the constancy of the speed of light—hereafter called the “light principle” for brevity—states that the speed of light is the same constant c in all inertial systems, independent of the relative velocity of the source and observer, and of direction, position, and time.³ We must inquire in what sense the null-result of the Michelson-Morley experiment can legitimately be held to support this relativistic light principle.

It will be evident from the results of our discussion that it is quite incorrect to suppose, as is done in some quarters, that the Michelson-Morley experiment furnishes a sufficient experimental basis for the light principle. And the statement of the *several* empirical and philosophical constituents of the light principle will then make apparent the *fallaciousness* of the following inference: since two different Galilean observers give formally the same description of the behavior of a single identical light pulse, emitted upon the coincidence of their origins, therefore the respective expanding spherical forms of this disturbance which they are said to “observe” in their respective systems are constituted by the selfsame point-events or the selfsame configuration of photons. Sometimes the ill-conceived attempt is made to render the false conclusion of this fallacious inference plausible by the irrelevant though true remark that while judgments can contradict one another, observed occurrences cannot. As if the

³ In opposition to the emission theory of light, the classical aether theory had already affirmed the independence of the velocity of light in the aether from that of its *source*, while affirming, however, its dependence on the observer's motion relative to the aether medium.

J. H. Rush has pointed out [*Scientific American*, Vol. CXCIII (August, 1955), p. 62] that the *time* constancy of the velocity of light may have to be questioned: measurements over the past century have yielded three fairly distinct sets of values for c . For a comprehensive recent survey of measurements of the velocity of light, see E. Bergstrand: “Determination of the Velocity of Light,” *Handbuch der Physik*, edited by Flügge (Berlin: J. Springer; 1956), Vol. XXIV, pp. 1–43.

claim that the light pulse spreads spherically in each inertial system were warranted by the deliverances of sensory observation alone and did not already presuppose a theory concerning the intersystemic and intrasystemic status of simultaneity.

In order to determine to what extent the light principle draws support from the outcome of the Michelson-Morley experiment, we must see first what logical grounds there are for not interpreting the null outcome of the Michelson-Morley experiment as evidence for the aether-theoretic Lorentz-Fitzgerald contraction hypothesis, an aether-theoretic auxiliary hypothesis to which we shall refer hereafter as the "L-F hypothesis" for brevity. The correct and detailed handling of this prior inquiry is particularly important, since the literature and the classroom instruction in the philosophy of science abound in the citation of the L-F hypothesis as an illustration of a logically *ad hoc* modification of a theory. And the purportedly *ad hoc* status of the L-F hypothesis is widely adduced as the justification for rejecting it as an interpretation of the null outcome of the Michelson-Morley experiment. I wish to show first that this characterization of the L-F hypothesis is as erroneous as it is hackneyed: the L-F hypothesis must be rejected because it is *false* and *not* because it is logically *ad hoc*. We shall also see that on the strength of the mistaken assessment of the aether-theoretic L-F hypothesis as logically *ad hoc*, the historical conclusion has been drawn that the theorizing of Lorentz and Fitzgerald involved a grave infraction of scientific method. My impending explanation of why it is that the L-F hypothesis cannot be regarded as logically *ad hoc* will therefore also demonstrate that philosophical misunderstandings unfoundedly generated a historical allegation concerning Lorentz and Fitzgerald.

An auxiliary hypothesis, introduced in response to new evidence embarrassing to an established theory, is *ad hoc* in the logical sense within the framework of the theory modified by it, if it does not, in principle, lend itself to any independent test whatever. Since the theoretical terms of the postulates of sophisticated empirical theories are given only a partial or incomplete physical interpretation by the rules of correspondence at any given time, these terms (and the concepts they represent) are "open" in the sense of admitting the adjunction of further rules

Correction on
pp. 722-724 of ch. 21
and in Append. §36

See Append. §35 for
amplification

of correspondence to the given theoretical postulates. Hence the class of all observational consequences of a physical theory is not a clearly circumscribed class, unless we specify and delimit the rules of correspondence which, in conjunction with the theoretical postulates, constitute the given established theory. Accordingly, if the question of whether a certain hypothesis auxiliary to an established theory is logically *ad hoc* is to have a well-defined status, it must be understood that this question is relativized to a specific, delimited set of rules of correspondence. In this way, the open-textured character of theoretical terms does not make for any imprecision in the property of being logically *ad hoc*. Now suppose that a certain collateral hypothesis such as the aether-theoretic L-F hypothesis is actually independently testable, at least in principle, but that its advocates fail to be aware of such testability and espouse it nonetheless. Clearly, the auxiliary hypothesis in question can then not be held to be logically *ad hoc* but can be regarded as *ad hoc* only in the psychological sense. For the methodological culpability of those who espouse such an auxiliary hypothesis despite their own (mistaken!) belief that it is not independently testable cannot detract from the actual, purely logical independent testability of the hypothesis in the context of its theoretical framework. Thus, we must distinguish between the logical and psychological senses of being *ad hoc*, much as we distinguish between the logical property possessed by a mathematical proposition that is a theorem within a given axiom system and the psychological attribute possessed by mathematicians who *realize* that the given proposition is indeed a theorem.

The proof which I am about to give that the original L-F hypothesis is not logically *ad hoc* in the context of the aether theory therefore need not take cognizance of the beliefs Lorentz and Fitzgerald actually entertained as to its independent testability.

A comparison of the reasoning underlying the Kennedy-Thorndike experiment⁴ with the design of the Michelson-Morley experiment will serve to establish that the aether-theoretic

See Append. §35 and
p. 718 of ch. 21

⁴ Cf. R. J. Kennedy and E. M. Thorndike: *Physical Review*, Vol. XLII (1932), p. 400, and W. Panofsky and M. Phillips: *Classical Electricity and Magnetism* (Cambridge, Mass.: Addison-Wesley Publishing Company, Inc.; 1955), p. 236.

L-F hypothesis did indeed have independently falsifiable consequences and that it was, in fact, falsified by the null outcome of the Kennedy-Thorndike experiment. Specifically, our analysis will show that: first, although the coupling of the L-F hypothesis with the classical aether theory entails the null outcome produced by the Michelson-Morley experiment, it does rule out the negative result that was actually yielded by the Kennedy-Thorndike experiment, and second, the aether theory as modified by the L-F hypothesis entails a positive outcome of the Kennedy-Thorndike experiment differing quantitatively from the positive result required by the aether theory without the L-F hypothesis.

The essential difference between the apparatus used in the Michelson-Morley and Kennedy-Thorndike experiments is the following: as measured by rods in the laboratory, the horizontal and vertical arms of the Michelson interferometer used in the Kennedy-Thorndike experiment are not equal but are made as different in length as possible, so as to assure a considerable difference in the travel times of the two partial beams from the source to the point at which they recombine to produce interference fringes. By contrast, the horizontal and vertical arms of the apparatus used in the Michelson-Morley experiment are each of the *same* length l as measured by the rods in the laboratory.

Accordingly, in the latter experiment the classically expected round-trip times T_v and T_h for the vertical and horizontal arms are given respectively by

$$T_v = \frac{2l}{(c^2 - v^2)^{\frac{1}{2}}} \text{ and } T_h = \frac{2l}{(c^2 - v^2)^{\frac{1}{2}}} \frac{1}{(1 - \beta^2)^{\frac{1}{2}}}$$

where c represents the velocity of light, v the velocity of the apparatus relative to the aether, and $\beta \equiv v/c$. Now, without a L-F hypothesis, the initial time difference $T_h - T_v$ between the two partial light beams would be expected to change in the course of the rotation of the apparatus through 90° in the Michelson-Morley experiment. And thus a shift in the interference fringes corresponding to this change was anticipated. But, once the aether theory is amended by the introduction of the L-F contraction, the length l in the expression for T_h must be replaced by the length $l(1 - \beta^2)^{\frac{1}{2}}$. As a consequence of the

introduction of this auxiliary hypothesis, T_h becomes *equal* to T_v , and the difference between the round-trip times of the two partial beams becomes zero throughout the Michelson-Morley experiment, in conformity with its null result.

It will be noted that the equality of the *terrestrially measured* lengths of the two arms is a necessary condition for the constant vanishing of the difference between the two round-trip times in the L-F account of the Michelson-Morley experiment. But precisely this necessary condition is not fulfilled in the Kennedy-Thorndike experiment, in which the terrestrially measured lengths of the vertical and horizontal arms have the unequal values L and l , respectively. Thus, *upon assuming the L-F hypothesis*, the difference between the travel times of the two light beams of the Kennedy-Thorndike experiment is given by

$$T_v - T_h = \frac{2L}{(c^2 - v^2)^{\frac{1}{2}}} - \frac{2l}{(c^2 - v^2)^{\frac{1}{2}}} = \frac{2}{(c^2 - v^2)^{\frac{1}{2}}} (L - l)$$

Instead of vanishing throughout the experiment, this time difference varies with the diurnally and annually changing velocity v of the apparatus relative to the fixed aether. Moreover, if we do not assume a L-F contraction, the difference between the two travel times of the Kennedy-Thorndike experiment has the *different* value given by

$$\begin{aligned} T_v - T_h &= \frac{2L}{(c^2 - v^2)^{\frac{1}{2}}} - \frac{2l}{(c^2 - v^2)^{\frac{1}{2}}} \cdot \frac{1}{(1 - \beta^2)^{\frac{1}{2}}} \\ &= \frac{2}{(c^2 - v^2)^{\frac{1}{2}}} \left(L - \frac{l}{(1 - \beta^2)^{\frac{1}{2}}} \right) \end{aligned}$$

And this difference is likewise a function of the diurnally and annually varying velocity v of the apparatus.

Under what conditions can the changing time difference called for by the L-F modification of the aether theory be expected to give rise to corresponding observable shifts in the interference fringe pattern? On that theory, these fringe shifts should occur if the clocks in the moving system are presumed to have the same rates as the "true" clocks of the aether system and if the frequency (period) of the light source—as measured by either the moving clocks or the aether-system clocks—does not itself depend upon the changing velocity of the apparatus with respect to the

aether. But since Kennedy and Thorndike found that the expected fringe shifts failed to materialize, the crucial question before us is whether the L-F hypothesis is falsified by this null result even though the positive result expected on the basis of the L-F form of the aether theory depended for its deduction also upon the assumption that the frequency of the light source is independent of its velocity through the aether. Hence we must inquire whether the null result of the Kennedy-Thorndike experiment could be explained by denying the latter assumption of independence while preserving the L-F hypothesis, so that this hypothesis would then not be falsifiable by this experiment.

Now, the denial of the assumption of independence could take either of the following two forms with a view to explaining the null result of Kennedy and Thorndike while affirming the L-F hypothesis:

a. The velocity-dependent time difference $T_v - T_h$ given above, which can be expressed as

$$T_v - T_h = \frac{2(L-1)}{c(1-\beta^2)^{\frac{1}{2}}},$$

does not give rise to fringe shifts, because the frequency of the moving source does depend on its velocity through the aether as follows: both its frequency as determined by the clocks of the aether system and its frequency as measured by the clocks of the moving system, the latter of which we shall call the "proper" frequency for brevity, are reduced by a factor of $(1-\beta^2)^{\frac{1}{2}}$ as compared with the frequency that would be measured in the aether system, if the moving source came to rest there. Accordingly, *an observational consequence of this way of denying the independence of the frequency from the velocity is that the proper frequency of the moving source would vary with its velocity v (and hence with β).*

b. The velocity-dependence of the time difference $T_v - T_h$ given under (a) does not issue in a fringe shift, because there is also the following compensatory velocity-dependence: both the frequency of the source, as measured by the aether-system clocks, *and* the rate of the clocks in the moving system, as compared with the clocks in the aether system, are reduced by a factor of $(1-\beta^2)^{\frac{1}{2}}$. Thus, the proper frequency is now independent of the

velocity of the source, and the values of the round-trip times as measured in the moving system are given by the quantities

$$T_v' = \frac{2L}{c} \quad \text{and} \quad T_h' = \frac{2l}{c}$$

which are independent of β .

We are now ready to state the grounds for a unique choice of the *third* from among the following three rival interpretations of the null result of the Kennedy-Thorndike experiment:

(1) The L-F hypothesis is confirmed *and* the observable *proper* frequency of the moving source *varies* with its velocity.

(2) The L-F contraction *and* the "time dilation" (or "dilatation," i.e., reduction of the rate of the moving clocks) operate together while the *proper* frequency of the source is *constant*, and hence the compound auxiliary hypothesis comprising *both* the L-F contraction *and* the time-dilation is confirmed.

(3) The L-F contraction hypothesis is *falsified* by the outcome of the Kennedy-Thorndike experiment in the sense of being (highly) disconfirmed. And since the Kennedy-Thorndike experiment constitutes a test which is independent of the Michelson-Morley experiment, the former experiment shows that the espousal of the L-F hypothesis in response to the negative result of the Michelson-Morley experiment was not logically *ad hoc*, although we shall see that it happened to have been psychologically *ad hoc*.

The first interpretation is to be rejected: its claim that the observable proper frequency of the moving source varies with the velocity is testable *apart* from the Kennedy-Thorndike experiment and has been found to be false by separate empirical evidence. The issue is therefore whether the fact that interpretation (2) is consistent with the null outcome of the Kennedy-Thorndike experiment renders (2) a *methodologically* acceptable interpretation and inductively a legitimate *alternative* to interpretation (3). That (2) does not, however, pass muster at all methodologically and hence cannot qualify as a rival to (3) emerges from the fact that the *combined* auxiliary hypothesis affirming *both* the L-F contraction *and* the time dilation is plainly *ad hoc*, since it does not lend itself to any independent test whatever. Hence, despite its compatibility with observational findings, methodo-

Correction in
Append. § 36

logical grounds prompt us to reject the doubly amended variant of the aether theory offered by (2). And it is now apparent that the attempt to endorse interpretation (2) in order to rescue the L-F hypothesis from the refutation that is claimed by interpretation (3) founders on the following fact: (2) succeeds in preserving the L-F hypothesis from refutation by the null outcome of the Kennedy-Thorndike experiment only by dint of incorporating that hypothesis in an augmented auxiliary hypothesis which is indeed logically *ad hoc*.

We have shown, therefore, that instead of being logically *ad hoc*, the L-F hypothesis has a falsifiable bearing on the outcome of the Kennedy-Thorndike experiment, an experiment which could have detected any existing velocity of the apparatus relative to the aether even on the assumption of a L-F contraction. And we note that the invalidation by the Kennedy-Thorndike experiment of the charge that the L-F hypothesis is logically *ad hoc* is *internal to the aether theory*: this invalidation no more depends logically on the availability of the STR than does the refutation of the original aether theory by the null outcome of the Michelson-Morley experiment. In addition to the Kennedy-Thorndike experiment, the following experiment suggested by C. Møller and others qualifies as a test of the L-F contraction within the framework of *prerelativistic* conceptions: although a *terrestrial* observer could not detect a L-F contraction in a Michelson-Morley experiment, "an observer at rest in the aether outside the earth would, however, in principle, be able to observe the shortening and he would find the earth and all objects on the earth contracted in the direction of motion of the earth."⁵

In view of the *falsity* of the supposition that the L-F hypothesis is logically *ad hoc*, that supposition cannot justifiably serve as the basis for the *historical* contention that Lorentz and Fitzgerald were unable to envision an independent test for their contraction hypothesis and hence were methodologically culpable for espousing it nonetheless.

But it is clear that laboring under the philosophical miscon-

⁵ C. Møller: *The Theory of Relativity*, *op. cit.*, p. 29. In Section F of the present chapter, we shall discuss the *important differences* between the *prerelativistic* and the *relativistic* conceptions of the status of any contraction disclosed by the extraterrestrial experiment suggested here by Møller.

ception that the L-F hypothesis is logically *ad hoc* could dissuade a historian of science from making the effort to uncover the kind of *historical* evidence which alone can show whether the L-F hypothesis was *psychologically ad hoc*.

H. Dingle⁶ has pointed out the existence of historical evidence establishing the biographical fact that, in the case of Lorentz, the contraction hypothesis was psychologically *ad hoc*. But unfortunately this historically documented fact then misled Dingle into inferring that the L-F hypothesis was logically *ad hoc*. As well conclude that a certain mathematical proposition cannot be a theorem in a given axiom system merely because all mathematicians are unaware that the proposition is, in fact, provable!

It is now clear that the null outcome of the Michelson-Morley experiment cannot be construed as support for the aether-theoretic L-F hypothesis, because there is indeed other evidence which shows that this hypothesis is false.

If then we therefore abandon the aether-theoretic framework of interpretation in favor of the relativistic one, we must return to the opening query of this Section and ask: to what extent can the relativistic light principle be claimed to rest on the null-outcome of the Michelson-Morley experiment?

The null-result of the Michelson-Morley experiment merely showed that if *within* an inertial system, light-rays are jointly emitted from a given point in different directions of the system and then reflected from mirrors at equal distances from that point, as measured by rigid rods, then they will return together to their common point of emission. And it must be remembered that the equality of the round-trip times of light in different directions of the system is measured here not, of course, by material clocks but by light itself (absence of an interference fringe shift).⁷ The repetition of this experiment at different times of the

⁶ H. Dingle: *The British Journal for the Philosophy of Science*, Vol. X (1959), pp. 228–29.

⁷ Since the classically expected time difference in the second-order terms is only of the order of 10^{-16} second, allowance must be made for the absence of a corresponding accuracy in the measurement of the equality of the two arms. This is made feasible by the fact that, on the aether theory, the effect of any discrepancy in the lengths of the two arms should *vary*, on account of the earth's motion, as the apparatus is rotated. For details, see P. Bergmann: *Introduction to the Theory of Relativity*, *op. cit.*, pp. 24–26, and J. Aharoni: *The Special Theory of Relativity* (Oxford: Oxford University Press; 1959), pp. 270–73.

year, i.e., in different inertial systems, showed only that there is no difference within *any* given inertial system in the round-trip times as between different directions within that system. But the outcome of the Michelson-Morley experiment does not show at all either that (a) the round-trip (or one-way) time required by light to traverse a closed (or open) path of length $2l$ (or l) has the same numerical value in different inertial systems, as measured by material clocks stationed in these systems⁸ or that (b) if, contrary to the arrangements in the Michelson-Morley and Kennedy-Thorndike experiments, the source of the light is outside the frame K in which its velocity is measured, then the velocity in K will be independent of the velocity of the source relative to K . In fact, the statement about the round-trip times made under (a) was first substantiated by the Kennedy-Thorndike experiment of 1932, as we shall see, and the assertion under (b) received its confirmation by observations on the light from double stars.⁹ Yet the light principle certainly affirms both (a) and (b) and thus clearly claims more than is vouchsafed by the Michelson-Morley experiment. Accordingly, we can see that in addition to the result of the Michelson-Morley experiment, the light principle contains at least all of the following theses:

(1) The assertion given under (a), which undoubtedly goes beyond the null-result of the Michelson-Morley experiment. For brevity, we shall call it "the clock axiom" in order to allude to its reference to clock-times of travel. Although lacking experimental corroboration at the time of Einstein's enunciation of the light principle, this "clock axiom" was suggested by the fundamental assumption of special relativity that there are no preferred inertial systems.

⁸ It is understood that the lengths $2l$ in the different frames are each ratios of the path to the same unit rod, which is transported from system to system in order to effect these measurements and which remains equal to unity *by definition* in the course of this transport.

⁹ Cf. J. Aharoni: *The Special Theory of Relativity*, *op. cit.*, pp. 269–70. The relativistic assumption stated under (b) was denied by Ritz's emission theory which maintained the following: the velocity of light in free space is always c with respect to the source, but its value in a frame K depends on the velocity of the source relative to K . For a careful account of the body of evidence against Ritz's hypothesis, see W. Pauli: *Theory of Relativity* (New York: Pergamon Press; 1958), pp. 6–9. Additional evidence against Ritz's assumption was furnished by a recent experiment by A. M. Bonch-Bruевич [*Physics Express*, November, 1960, pp. 11–12].

(2) The Einstein postulate regarding the maximal character of the velocity of light, which is a source of the conventionality and relativity of simultaneity within a given inertial system and as between different systems respectively. This postulate *allows* but does *not* entail that we choose the same value $\epsilon = 1/2$ for all directions within any given system and also for each different system.

(3) The claim that the velocity of light in any inertial system is independent of the velocity of its source.

To see specifically how the light principle depends for its validity upon the truth of all of these constituent theses, we now consider the consequences of abandoning any one of them while preserving the remaining two. To do so, we direct our attention to the one-way velocity of an outgoing light pulse that traverses a distance l in system S and also to the one-way velocity of such a pulse traversing a distance l in system S' . Let T_s and $T_{s'}$ be the respective outgoing transit times of these light pulses for the distances l in S and S' , the corresponding round-trip times in these frames being τ_s and $\tau_{s'}$. In this notation, our earlier equation $t_2 = t_1 + \epsilon(t_3 - t_1)$ of Section B becomes

$$T_s = \epsilon \tau_s,$$

and similarly for S' . Hence the relevant one-way velocities of light in these systems are, respectively, given by

$$v_s = (l/T_s) = (l/\epsilon \tau_s)$$

and

$$v_{s'} = (l/\epsilon \tau_{s'}).$$

Our problem is to determine, one by one, the consequences of assuming that of the three above ingredients of the light principle, only two can be invoked at a time in an attempt to assure that if $v_s = c$ in virtue of the observed value of τ_s and of a choice of $\epsilon = 1/2$ for S , then $v_{s'}$ will also have the value c .

(1) In the absence of the clock axiom, it can well be that $\tau_s \neq \tau_{s'}$. In that case, the relations $v_s = (l/\frac{1}{2}\tau_s) = c$ and $v_{s'} = (l/\epsilon \tau_{s'})$ tell us that it will not be possible to make $v_{s'} = c$ by a choice of $\epsilon = \frac{1}{2}$ in S' . To be sure, it may still be possible then to choose an appropriately different value of ϵ for S' so as to make $v_{s'} = c$. But, much as the latter choice of ϵ might yield

the value c for the outgoing one-way velocity of light in S' , it would also inescapably entail a value different from c for the return-velocity. And such a result would not be in keeping with the light principle.

(2) If, however, we guarantee that $\tau_{*} = \tau_{*}'$ by assuming the clock axiom but disallow the freedom to choose a value of ϵ for each inertial system by withdrawing the second constituent thesis above, then v_{*}' could readily be different from c . For the illegitimacy of choosing ϵ to be $\frac{1}{2}$ in S' would then derive from physical facts incompatible with the conventionality of simultaneity which would objectively fix the value of T_{*}' as different from $\frac{1}{2}\tau_{*}'$.

(3) The need for the third ingredient is obvious in view of what has already been said.

It is of importance to note that even when all of the above constituent principles of the light principle are assumed, no fact of nature independent of our descriptive conventions would be contradicted, if we chose values of ϵ other than $\frac{1}{2}$ for each inertial system, thereby making the velocity of light different from c in both senses along each direction in all inertial systems. I have shown elsewhere¹ that this conclusion is not invalidated either by determinations of the one-way velocity of light on the basis of measurements of the half-wave-length of standing waves in cavity resonators or by measurements of the ratio of the electromagnetic and electrostatic units of charge. The assertion that the invariant velocity of light is c is therefore not, in its entirety, a purely factual assertion. But we saw that it is nonetheless a consequence of presumed physical facts that the specification of the velocity of light involves a stipulational element which, in combination with other factual principles, allows us to say that the velocity of light is c .

(E) THE EXPERIMENTAL CONFIRMATION OF THE KINEMATICS OF THE STR.

We saw that the Michelson-Morley experiment clearly could not be regarded as empirical proof for the "clock axiom" ingredient in the light principle. It is therefore a welcome fact that the

¹ A. Grünbaum: *American Journal of Physics*, Vol. XXIV (1956), pp. 588-90.

Kennedy-Thorndike experiment described in Section D of this chapter can indeed be invoked as support for that axiom's relativistic denial of the existence of preferred inertial systems.²

If the Kennedy-Thorndike experiment had yielded a positive effect instead of the null-result which it did actually yield, then it could have been cogently argued that the Michelson-Morley experiment was evidence for a bona fide Lorentz-Fitzgerald contraction, just as a fringe shift produced by heating one of the arms of the interferometer could be held to be evidence for the elongation of that arm. But, in view of the de facto null outcome of the Kennedy-Thorndike experiment, there is very good reason indeed to attribute the absence of a diurnal or annual variation in the time-difference between the two partial beams to a *constancy, as between different inertial systems, of the time required by each of the partial beams to traverse its own closed path*. And thus we are entitled to say that the Kennedy-Thorndike experiment has provided empirical sanction for the clock axiom.³

² J. P. Cedarholm and C. H. Townes have reported their very recent cognate experiment [cf. "A New Experimental Test of Special Relativity," *Nature*, Vol. CLXXXIV (No. 4696, 1959), pp. 1350-51]. They write:

The experiment compares the frequencies of two maser oscillators with their beams of ammonia molecules pointed in opposite directions, but both parallel to a supposed direction of motion through the aether. . . . A precision of one part in 10^{12} has been achieved in this frequency comparison, and failure to find a frequency change of the predicted type allows setting the upper limit on an aether drift as low as 1/1000 of the orbital velocity of the Earth. . . . The present experiment sets an upper limit on an aether-drift velocity about one-fiftieth that allowed by previous experiments.

³ It should be mentioned that the experimenters Kennedy and Thorndike themselves conceived of their experiment not as a test of what I have called the "clock axiom" in the light principle but rather as a test of the relativistic clock retardation. The negative outcome of the Kennedy-Thorndike experiment cannot, however, be invoked as evidence for the *relativistic* clock retardation, which is a reciprocal or symmetrical relation between the clocks of relatively moving inertial systems. For the reader will recall my account in Section D of the reasons for *rejecting* the interpretation that the null result of the Kennedy-Thorndike experiment supports the hypothesis that the *aether-theoretic* time-dilation and the L-F contraction operate together. And it is evident from that analysis in Section D that instead of pertaining to the reciprocal *relativistic* clock retardation, the null-outcome of the Kennedy-Thorndike experiment has a bearing on the *ad hoc* assumption of an *aether-theoretic* time dilation, which is *not* a symmetrical or reciprocal relation between the clocks of the moving system and those of the aether-system.

We recall that in 1905, there was no unambiguous experimental evidence supporting Einstein's postulate that light is the fastest signal. But subsequent experiments showed that the mass and kinetic energy of accelerated particles become indefinitely large as their velocity approaches that of light.⁴

There is therefore impressive empirical evidence for all of the constituent theses of the light principle. It would be an error, however, to suppose that the experimental justification of the light principle suffices also to substantiate the Lorentz transformations. For these equations entail the clock retardation, whereas the light principle alone does not. Hence it is a mistake to suppose that upon a suitable choice of zeros of time, the light principle alone permits the deduction of the Lorentz transformations and that all the novel affirmations of relativistic kinematics are thus vouchsafed by the light principle. A lucid demonstration of the fact that the relativistic clock retardation is logically independent of the light principle has recently been given by H. P. Robertson.⁵ He considers a linear transformation

$$T: (t', x', y', z') \rightarrow (t, x, y, z)$$

between a kind of primary inertial system Σ and a "moving" inertial system S' . Upon having fixed thirteen of the sixteen coefficients of this transformation by various conventions, symmetry conditions, and a specification of the velocity v of S' relative to Σ , Robertson is concerned with the experimental warrant for asserting that the remaining three coefficients have the values required by the Lorentz transformations. And he then shows that the Michelson-Morley and Kennedy-Thorndike experiments, which do succeed in completing the confirmation of the light principle, do not suffice to fix the remaining three coefficients of the transformation such that these have the values required by

⁴ See W. Gerlach: *Handbuch der Physik* (Berlin: Springer-Verlag; 1926), pp. 61 ff., and C. Møller: *The Theory of Relativity*, *op. cit.*, Chapter iii, Sec. 32.

⁵ H. P. Robertson: *Reviews of Modern Physics*, Vol. XXI (1949), p. 378. An earlier proof of the compatibility of the light principle with the denial of the clock retardation was given by Reichenbach (see *Axiomatik der relativistischen Raum-Zeit-Lehre*, *op. cit.*, pp. 79–83, esp. pp. 81–83), who exhibits a consistent set of coordinate transformations embodying both assertions.

the Lorentz transformations. An additional experiment is needed to do so: the laboratory work of Ives and Stilwell (1938) furnished the lacking data by observations on high speed canal rays. And it was their confirmation of the relativistic ("quadratic") Doppler effect⁶ that constituted the first experimental proof of the clock retardation affirmed by the Lorentz transformations. Additional confirmation has been provided by data on the rate of disintegration of mesons.⁷

According to the STR, the clock retardation is exhibited by all natural clocks, be they the material clocks which keep astronomical time or atomic clocks like cesium atoms, whose unit the theory assumed to have a time-invariant ratio to the astronomical unit. But this assumption of a constant ratio of the atomic and astronomical units of time has been questioned by Dirac, Milne, Jordan, and others, who have suggested that the ratio increases continuously by amounts of the order of $1/T$ per year, where T is "the age of the universe" in years. If T is about 4×10^9 years, the change at the present epoch is of the order of $1/10^9$. Changes of this order may soon be measurable over intervals of a few years, since Essen and Parry of the National Physical Laboratory in England have recently measured the natural resonant frequency of the cesium atom with a stated precision of 1 in 10^9 and hope for even higher accuracies.⁸

(F) THE PHILOSOPHICAL ISSUE BETWEEN EINSTEIN AND HIS AETHER-THEORETIC PRECURSORS, AND ITS BEARING ON E. T. WHITTAKER'S HISTORY OF THE STR.

In view of our discussion of Newton's conception of the metrics of space and time in Chapter One, it will suffice at this point to say quite briefly that the aether-theoretic version of the New-

⁶ Cf. M. von Laue: *Die Relativitätstheorie*, *op. cit.*, p. 20.

⁷ B. Rossi and D. B. Hall: *Physical Review*, Vol. LIX (1941), p. 223. For a proof that in the context of the conditions to which the STR is applicable, the relativistic clock retardation does not give rise to the "clock paradox," see A. Grünbaum: "The Clock Paradox in the STR," *Philosophy of Science*, Vol. XXI (1954), pp. 249-53 and Vol. XXII (1955), pp. 53 and 233.

⁸ Cf. L. Essen and J. V. L. Parry: *Nature*, Vol. CLXXVI (1955), p. 280 and the comments by G. M. Clemence: *ibid.*, p. 1230, and in *Science*, Vol. CXXIII (1956), p. 571. See also H. Lyons: *Scientific American*, Vol. CXCVI (February, 1957), p. 71.

tonian theory included the following thesis concerning the absolute rest system of the aether: receptacle space and time each have their own *intrinsic* metric, which exists quite independently of the presence of material rods and clocks in the universe, devices whose function is *at best* the purely epistemic one of enabling us to ascertain the intrinsic metrical relations of the receptacle space and time contingently containing them. Thus, for example, even when material clocks run uniformly, they are merely in accord with but do not "define" the temporal metric.

Awareness of this thesis of the aether-theory will enable us shortly to give a precise statement of the fundamental *logical* differences between the aether-theoretic L-F contraction and the *numerically* equal relativistic "Einstein contraction" of a moving rod, which is deducible from the relativistically interpreted Lorentz transformations. A clear delineation of the logical differences between the L-F and Einstein contractions is especially important since these contractions are often confounded. Thus, we are told that instead of attempting to account for the null result of the Michelson-Morley experiment via the purportedly *ad hoc* L-F contraction hypothesis, Einstein *explained* the outcome of that experiment soundly by deducing the contraction from the Lorentz transformations. To cite a very recent case in point, R. M. Palter, writing on the kinematic results of the STR, credits Einstein with explaining the outcome of the Michelson-Morley experiment by saying: "the contraction of rigid rods was, of course, confirmed by the Michelson-Morley and numerous subsequent aether-drift experiments."⁹ But, as is already evident from our account of the constituent theses of the light principle in Section D, far from explaining the outcome of the Michelson-Morley experiment non-trivially as a consequence of more fundamental principles, Einstein incorporated its null result as a physical *axiom* in his light principle, which is a *premise* in his deduction of the Lorentz transformations.

More fundamentally, the erroneous view that the relativistic deduction of the Einstein contraction from the relativity of simultaneity ingredient in the Lorentz transformations *explains* the null-outcome of the Michelson-Morley experiment derives its

⁹ R. M. Palter: *Whitehead's Philosophy of Science, op. cit.*, p. 13.

See Append. §37

plausibility from the numerical equality of the contraction factors of the relativistic Einstein contraction and the aether-theoretic L-F contraction, the value being $\sqrt{1-\beta^2}$ in each case. But these contractions are of radically different logical character, because the L-F contraction pertains to the very system in which the contracted arm is *at rest*, whereas the contraction that Einstein derived from the Lorentz transformations pertains to the length measured in a system relative to which the arm is *in motion*. More explicitly, the Lorentz-Fitzgerald contraction hypothesis asserts a comparison of the actual length of the arm, as measured by the round-trip time of light, to the greater length that the travel-time would have revealed, if the classical aether theory were true. Thus, using light as the standard for effecting the comparison, this hypothesis affirms that in the same system and under the same conditions of measurement, the metrical properties of the arm are different from the ones predicted by classical aether theory. And this difference or contraction is clearly quite independent of any contraction based on comparisons of lengths in different inertial systems. By contrast, the contraction which Einstein deduced from the Lorentz transformations is based on a comparison of the length of a rod, as measured from an inertial system relative to which it is in motion, to the length of that same rod, as measured in its own rest system. Unlike the L-F contraction, but like the relativistic clock retardation, this "Einstein contraction" is a *symmetrical* or reciprocal relation between the measurements made in any two inertial systems and is a consequence of the intersystemic relativity of simultaneity, because it relates lengths determined from *different* inertial perspectives of measurement, instead of contrasting conflicting claims concerning the results obtained under the same conditions of measurement. What Einstein did explain, therefore, is this "metrogenic" contraction, a phenomenon which poses no greater logical difficulties than the differences in the angular sizes of bodies that are observed from different distances.

Lest it be thought that the confusion of the two kinds of contraction just discussed is a thing of the past, I cite the following recent statements from Sir E. T. Whittaker's *Turner Lectures*:

The Lorentz transformation . . . supplies at once an explanation of the Fitzgerald contraction. . . . the failure of all attempts

to determine the velocity of the earth by comparing the Fitzgerald contraction in rods directed parallel and perpendicular to the terrestrial motion . . . is necessitated by the Postulate of Relativity. But there is no impossibility in principle at any rate, in observing the contraction, provided we can make use of an observation-post which is *outside* the moving system.¹

To be sure, if Whittaker's proviso of an aether-system observation post outside the moving earth be granted, then an observer at that post who interprets his data *pre-relativistically* would confirm the Lorentz-Fitzgerald contraction by finding that the "true" length of the moving arm, which he believes himself to be observing from his vantage point, is smaller than the "spurious" length measured by the rod of a terrestrial observer. But the relativistic explanation of the numerically equal Einstein contraction actually involved here *rejects*, as we shall see in detail, the very conceptions which alone give meaning to (a) construing the findings of the extraterrestrial observer as *equivalent* to a contraction *within* the moving system in the sense of Lorentz and Fitzgerald, and (b) asking within the framework of the STR in the spirit of E. T. Whittaker why there is a L-F contraction within the moving system. And the relativistic deduction of the Einstein contraction from the Lorentz transformations can therefore have no bearing at all in Whittaker's sense on why the Michelson-Morley experiment failed to fulfill the predictions of the classical aether theory.

We see that the locus of the philosophical difference between Lorentz and Einstein has been misplaced by those who point to Einstein's alleged explanation of the null result of the Michelson-Morley experiment. To understand Einstein's philosophical innovation, we must take cognizance of the fact that the L-F contraction hypothesis was not the only addition to the aether theory made by Lorentz in order to account for the available body of experimental data. In addition, he had been driven to postulate with J. Larmor that just as rods are caused to contract in any inertial system moving relatively to the aether, so also clocks are

¹ E. T. Whittaker: *From Euclid to Eddington* (London: Cambridge University Press; 1949), pp. 63-64; italics supplied. The same error is repeated by Whittaker in his *A History of the Theories of Aether and Electricity*, *op. cit.*, Vol. II, p. 37, which should be read in conjunction with his "G. F. Fitzgerald," *Scientific American*, Vol. CLXXXIX (November, 1953), p. 98.

caused by that very motion to modify their rates and to read a spurious "local" time (as distinct from the "true" time shown by the clocks in the aether-system). The conceptual framework of Lorentz's interpretation of the transformation equations known by his name was the absolutistic one in which the clocks and rods even in the privileged aether system merely *accorded with* but did not "define" the metric of container space and time. On this basis, Lorentz was led to reason somewhat as follows:

(1) Since the horizontal arm of the Michelson-Morley experiment is shorter than a rod lying alongside it but conforming to the expectations of classical optics and its container theory of space, we must infer that when a unit rod in the aether system is transported to a moving system, it can no longer be a true unit rod but becomes shorter than unity in the moving system, and similarly for clocks.

(2) The deviation from the classically expected behavior exhibited by rods and clocks must have a cause in the sense of being due to a perturbational influence. For in the absence of such a cause, the classically expected behavior would have occurred spontaneously.

Einstein left the Lorentz transformations formally unaltered. But the reasoning underlying his radical reinterpretation of their physical meaning was based on the relational conception of space and time discussed in Chapter One. On the relational conception of space, the length of a body AB is an attribute of the relation between two pairs of points: the termini of AB on the one hand and those of the chosen unit rod on the other. Similarly the length of a time-interval is a ratio of that interval to the unit defined by some periodic physical process. And it is inherent in this definition of length as a ratio that the unit rod be at rest relative to AB when performing its metrical function.² But, on this relational

² It is perfectly clear that relations or relational properties of physical objects (which are expressed numerically as ratios) are fully as objective physically and exist just as independently of the human mind as simple properties of individual objects. Thus, the relation between a copper bar at rest in a system K and the unit rod in K might have the property that the copper bar has a length of 5 units in K. But the different relation of that bar's projection onto the x-axis of a system S, relative to which it is moving along that axis, to the unit rod of S may then yield a length of only

conception of length, one is not entitled to infer with Lorentz that a unit rod in the aether system will no longer be unity, *as a matter of physical fact*, once it has been transported to a moving system, and that such transport renders spurious the "local" readings of clocks in moving systems. For the relational conception allows us to call that same rod unity in the moving system by *definition*, and similarly for the units of time on clocks at rest in a moving system. If Lorentz had realized that the length of this rod in the moving system can be legitimately *decreed* by definition, and similarly for the periods of material clocks, then it would have been clear to him that the ground is cut from under his distinction between "true" ("real") and "local" (i.e., spurious or apparent) lengths and times and thereby from his idea that the horizontal arm in the Michelson-Morley experiment is actually shorter than the vertical arm. Thus, a coupling of the epistemological insights of the relational theory of length with the experimental findings of optics deprives reference to a preferred

4.7 S-units. It is incontestable that the differences among the various relations sustained by the bar do *not* render these relations subjective products of the physicist's mind, any more than they do the fact that the bar in question is a copper bar. In a futile attempt to defend a mentalistic metaphysics on the basis of relativity theory, Herbert Dingle denies this fact. Replying to decisive critiques of his views by P. Epstein (*American Journal of Physics*, Vol. X [1942], pp. 1 and 205, and Vol. XI [1943], p. 228), and M. Born (*Philosophical Quarterly*, Vol. III [1953], p. 139), Dingle offers the following argument (*The Sources of Eddington's Philosophy* [London: Cambridge University Press; 1954], pp. 11-12):

The view that physics is the description of the character of an independent external world was simply no longer tenable. . . . Every relativist will admit that if two rods, A and B, of equal length when relatively at rest, are in relative motion along their common direction, then A is longer or shorter than B, or equal to it, exactly as you please. It is therefore impossible to evade the conclusion that its length is not a property of either rod; and what is true of length is true of every other so-called physical property. Physics is therefore (*sic!*) not the investigation of the nature of the external world.

Far from having demonstrated that relativity physics is subjective, Professor Dingle has merely succeeded in exhibiting his unawareness of the fact that *relational* properties do not cease to be bona fide objective properties just because they involve relations between individuals rather than belong directly to individuals themselves. Only such unawareness can lead to his primitive thesis that the relations of physical entities to one another cannot constitute "the character of an independent external world."

aether system of all objective physical significance and makes possible the enunciation of the principle of relativity.

We can understand Einstein's philosophical departure from the second step in Lorentz's reasoning by giving an analysis of Lorentz's invocation of a cause for the contraction revealed by the Michelson-Morley experiment.

Every physical theory tells us what particular behavior of physical entities or systems it regards as "natural" in the absence of the kinds of *perturbational* influences which it envisions. Concurrently, it specifies the influences or causes which it regards as responsible for any deviations from the assumedly "natural" behavior. But when such deviations are observed and a theory cannot designate the perturbations to which it proposes to attribute them, its assumptions concerning the character of the "natural" or unperturbed behavior become subject to doubt. For the reliability of our conceptions as to what pattern of occurrences is "natural" is no greater than the *scope* of the evidence on which they rest. And a theory's failure to designate the perturbing causes of the nonfulfillment of its expectations therefore demands the envisionment of the possibility that: first, the "natural" behavior of things is indeed different from what the theory in question has been supposing it to be and that, second, deviations from the assumedly natural behavior transpire *without* perturbational causes of the kind previously envisioned by the theory.

Several examples from past and present scientific controversy attest to the mistaken search for the previously envisioned kinds of perturbational factors which are supposed to cause deviations from the pattern which a particular theory unquestioningly and tenaciously affirms to be the natural one. Thus, Aristotelian critics of Galileo, assuming that Aristotle's mechanics describes the natural behavior, asked Galileo to specify the cause which prevents a body from coming to rest and maintains its speed in the same straight line in the manner of Newton's first law of motion. It was axiomatic for them that uniform motion could not continue indefinitely in the absence of net external forces. In our own time, there are those who ask: if the "new cosmology" of Bondi and Gold is true, must there not be a divine interference (perturbation) in the natural order whereby the spontaneous accre-

tion ("creation") of matter is caused?³ The propounders of this question do not tell us, however, on what basis they take it for granted that the total absence of any accretion of matter, however small or slow, is *cosmically* the natural state of affairs. Curiously enough, some theologians ask us to regard a state of "nothingness" (whatever that is!) as the natural state of the universe and thereby endeavor to create grounds for arguing that the mere existence and conservation of matter or energy require a divine creator and sustainer. But suppose that it were indeed to turn out that the physical world is a three-dimensional expanding spherical space whose radius of curvature had a minimum value at one time and that the latter was a genuine beginning because it was not preceded by an earlier contraction of which the present expansion is an "elastic" rebound. Then within the framework of that theory of the expanding universe, *every* phase of this expansion process must be held to be integral to the *natural* behavior of the universe. In that case, the inception of the expansion does not constitute evidence for the operation of a divine creator as perturbational cause, any more than any other phase of the expansion does. For there is absolutely no logically viable criterion for distinguishing those newly observed or inferred facts which merely compel the revision of previous generalizations and which the proponent of divine creation regards as part of the "natural order" from those to which he gratuitously attributes the status of being "outside" the "natural order" and of therefore being due to divine intervention in that order. We see that the theological concept of miracles attributes a supernatural origin to certain phenomena by the gratuitous declaration that a certain set of limited empirical generalizations which do not allow for these phenomena define with *certainty* what is natural. As if observed events presented themselves to us with identification tags as to their "naturalness"!⁴

³ The logical blunder which generates this question was committed in inverted form by Herbert Dingle. One of his reasons for rejecting the cosmology of Bondi and Gold is that it would allegedly require not merely a single act of miraculous divine interference, as Biblical creation *ex nihilo* does, but a continuous series of such acts. See A. Grünbaum: *Scientific American*, Vol. CLXXXIX (December, 1953), pp. 6-8.

⁴ For further details on this issue, see A. Grünbaum: *Scientific Monthly*, Vol. LXXIX (1954), pp. 15-16.

The basis for Einstein's philosophical objection to the second step in Lorentz's reasoning is now at hand: it was an error on Lorentz's part to persist, in the face of mounting contrary evidence, in regarding the classically expected behavior as the natural behavior. It was this persistence which forced him to explain the observed deviations from the classical laws by postulating the operation of a physically non-designatable aether as a perturbational cause. Having used the relational theory of length and time to reject the conclusion of the first step in Lorentz's reasoning, Einstein was able to see that the unexpected results of the Michelson-Morley experiment do not require any perturbational causes of the kind envisioned in the aether theory because they are integral to the "natural" behavior of things. This is not to say that what is taken axiomatically as the "natural" or unperturbed behavior of particles and light within the framework of the STR might not be regarded as in need of explanation by reference to perturbing causes within the context of a wider and/or deeper theory such as the GTR.

The character and significance of these fundamental philosophical differences between Einstein's conception of the Lorentz transformations, on the one hand, and the interpretation of earlier versions of these transformations by his predecessors Fitzgerald, Larmor, Poincaré, and Lorentz, on the other, altogether escaped recognition by E. T. Whittaker in his very recent account of the history of the development of the STR. His account of the history of the STR in his monumental *History of the Theories of Aether and Electricity* provides a telling illustration of how philosophical misconceptions can issue in a grievously false historical estimate of the relative contributions of a conceptual innovator vis-à-vis those of his precursors. In particular, we shall now see that a correct philosophical construal of how Einstein understood the space and time coordinates of the Lorentz transformations could have precluded Whittaker's unsound assessment of the contributions made to the STR by Lorentz, Poincaré, Fitzgerald, Larmor, and Voigt, an assessment which lent credence to his disparaging evaluation of Einstein's role.

Entitling his chapter on the history of the development of the STR, "The Relativity Theory of Poincaré and Lorentz," Whittaker gives the following depreciatory evaluation of Einstein's

role in the genesis of the STR: "In the autumn of the same year [1905] . . . Einstein published a paper which set forth the relativity theory of Poincaré and Lorentz with some amplifications, and which attracted much attention. . . . In this paper Einstein gave the modifications which must now be introduced into the formulas for aberration and the Doppler effect."⁵ This historical assessment of the range of Einstein's originality and of the magnitude of his contribution to the STR derives at least its plausibility (even if perhaps not its inspiration) from Whittaker's *philosophically incorrect* conception of *Einstein's interpretation* of the Lorentz transformations. For Whittaker committed two grievous logical errors. He failed to be cognizant of Einstein's philosophical *repudiation* of the Lorentz-Larmor-Poincaré distinction between "true" or "real" and "spurious" (apparent, local) times and lengths, a repudiation which renders Einstein an authentic conceptual revolutionary of genius in this respect. And correlatively, as we saw earlier, he made no allowance at all for the crucial *logical* differences between the pre-Einsteinian and Einsteinian conceptions of the status of any extraterrestrially observed contraction of a rod moving with the earth and pointing in the direction of the latter's motion.

Thus, only philosophical awareness of the fact that Einstein's conception of the Lorentz transformations is not to be construed along the modified aether-theoretic lines of Lorentz and Poincaré makes possible the discernment of the mistake in Whittaker's historical treatment of the STR.

⁵ E. T. Whittaker: *A History of the Theories of Aether and Electricity*, *op. cit.*, Vol. II, p. 40. Whittaker himself points out (p. 36) that even at the time of his death in 1928, Lorentz reportedly still favored the concepts of "true" time and absolute simultaneity. For Lorentz's own brief statement on this point, see his *The Theory of Electrons* (New York: Columbia University Press; 1909), pp. 329–30. In regard to the role played by Lorentz and Poincaré as precursors of Einstein, von Laue gives evidence in a more recent publication (*Naturwissenschaften*, Vol. XLIII [1956], p. 4) that Einstein was unaware of the groundwork they had done for the theory of relativity.

Chapter 13

PHILOSOPHICAL APPRAISAL OF E. A. MILNE'S ALTERNATIVE TO EINSTEIN'S STR

E. A. Milne, whose two logarithmically related t and τ scales of time were mentioned in Chapter One, Section C, has attempted to erect the usual space-time structure of the STR on the basis of a light signal kinematics of particle observers purportedly dispensing with the use of rigid solids and isochronous material clocks.¹ In his *Modern Cosmology and the Christian Idea of God*,² Milne begins his discussion of time and space by *incorrectly* charging Einstein with failure to realize that the concept of a rigid body as a body whose rest length is invariant under transport contains a conventional ingredient just as much as does the concept of metrical simultaneity at a distance.³ Milne then proposes to improve upon a rigid body criterion of spatial congruence by proceeding in the manner of radar ranging and using instead the round-trip times required by light to traverse the corresponding closed paths, these times not being measured by material clocks but, in outline, as follows.⁴ Each particle is equipped with

¹ E. A. Milne: *Kinematic Relativity*, *op. cit.*, and *Modern Cosmology and the Christian Idea of God*, *op. cit.*

² *Ibid.*, Chapter iii.

³ That Einstein was abundantly aware of this point is evident from his definition of the "practically rigid body" in *Geometrie und Erfahrung*, *op. cit.*, p. 9.

⁴ For more detailed summaries of Milne's light-signal kinematics cf. A. G. Walker: "Axioms for Cosmology" in L. Henkin, P. Suppes, and A. Tarski

a device for ordering the genidentical events belonging to it temporally in a linear Cantorean continuum. Such a device is called a "clock," and the single observer at the particle using such a local clock is called a "particle-observer." If now A and B are two particle observers and light signals are sent from one to the other, then the time of arrival $\bar{t}$ at B can be expressed as a function $\bar{t} = f(t)$ of the time t of emission at A, and likewise the time of arrival t' at A is a function $t' = F(\bar{t})$ of the time $\bar{t}$ of emission at B. Particle-observers equipped with clocks as defined are said to be "equivalent," if the so-called signal functions f and F are the same, and the clocks of equivalent particle-observers are said to be *congruent*. It can be shown that if A and B are not equivalent, then B's clock can be regraduated by a transformation of the form $\bar{t}' = \Psi(\bar{t})$ so as to render them equivalent.⁵ The congruence of the clocks at A and B does not, of course, assure their synchronism. Milne now uses Einstein's definition of simultaneity:⁶ the time t_2 assigned by A to the arrival of a light ray at B which is emitted at time t_1 at A and returns to A at time t_3 after instantaneous reflection at B is defined to be

$$t_2 = \frac{1}{2}(t_1 + t_3).$$

And he defines the distance r_2 of B, *by A's clock, upon the arrival of the light from A at B* to be given by the relation

$$r_2 = \frac{1}{2} c (t_3 - t_1),$$

where c is an *arbitrarily* chosen constant.⁷ Since

$$\frac{r_2}{t_2 - t_1} = \frac{r_2}{t_3 - t_2} = c,$$

the constant c represents the velocity of the light-signal in terms of the conventions adopted by A for measuring distance and time at a remote point B. Milne gives the following statement of his epistemological objections to Einstein's use of rigid rods and

(eds.) *The Axiomatic Method* (Amsterdam: North Holland Publishing Company; 1959), pp. 309-10, and L. Page and N. I. Adams: *Electrodynamics* (New York: D. Van Nostrand Company; 1940), pp. 78-85.

⁵ E. A. Milne: *Modern Cosmology and the Christian Idea of God*, *op. cit.*, pp. 39-41.

⁶ *Ibid.*, p. 42.

⁷ *Ibid.*

of his claim that his light-signal kinematics provides a philosophically satisfactory alternative to it:

the concept of the transport of a rigid body or rigid length measure is itself an indefinable concept. In terms of one given standard metre, we cannot say what we mean by asking that a given "rigid" length measure shall remain "unaltered in length" when we move it from one place to another; for we have no standard of length at the new place. Again, we should have to specify standards of "rest" everywhere, for it is not clear without consideration that the "length" will be the same, even at the same place, for different velocities. The fact is that to say of a body or measuring-rod that it is "rigid" is no definition whatever; it specifies no "operational" procedure for testing whether a given length-measure after transport or after change of velocity is the same as it was before. . . .⁸

It is part of the debt we owe to Einstein to recognize that only "operational" definitions are of any significance in science . . . Einstein carried out his own procedure completely when he analysed the previously undefined concept of simultaneity, replacing it by tests using the measurements which have actually to be employed to recognize whether two distant events are or are not simultaneous. But he abandoned his own procedure when he retained the indefinable concept of the length of a "rigid" body, i.e., a length unaltered under transport. The two indefinable concepts of the transportable rigid body and of simultaneity are on exactly the same footing; they are fog-centres, inhibiting further vision, until analysed and shown to be equivalent to conventions. . . .⁹

It will be one of our major tasks to elucidate the type of graduation employed for graduating our ordinary clocks; that is to say, to inquire what is meant by, and if possible to isolate what is usually understood by, "uniform time." In other words, we wish to inquire which of the arbitrarily many ways in which the markings of our abstract clock may be graduated can be identified with the "uniform time" of physics. . . .¹

The question now arises: is it possible to arrange that the mode of graduation of observer B's clock corresponds to the

⁸ *Ibid.*, p. 35.

⁹ *Ibid.*

¹ *Ibid.*, p. 37.

mode of graduation of A's clock in such a way that a meaning can be attached to saying that B's clock is a copy of A's clock? If so, we shall say that B's clock has been made congruent with A's. . . .²

It will have been noticed that we have succeeded in making B's clock a copy of A's without bringing B into permanent coincidence with A. We have made a copy of an arbitrary clock *at a distance*. This is something we cannot do with metre-scales or other length-measures. The problem of copying a clock is in principle simpler than the problem of copying a unit of length. We shall see in due course that with the construction of a copy of a clock at a distance we have solved the problem of comparing lengths. . . .³

The important point is that epoch and distance (which we shall call coordinates) are purely conventional constructs, and have meaning only in relation to a particular form of clock graduation. . . . But it is to be pointed out that when the mode of clock graduation reduces to that of ordinary clocks in physical laboratories, our coordinate conventions provide measures of epoch and distance which coincide with those based on the standard metre. . . .⁴

The reason why it is more fundamental to use clocks alone rather than both clocks and scales or than scales alone is that the concept of the clock is more elementary than the concept of the scale. The concept of the clock is connected with the concept of "two times at the same place," whilst the concept of the scale is connected with the concept of "two places at the same time." But the concept of "two places at the same time" involves a convention of simultaneity, namely, simultaneous events at the two places, but the concept of "two times at the same place" involves no convention; it involves only the existence of an *ego*. . . .⁵

Length is just as much a conventional matter as an epoch at a distance. Thus the metre-scale is not such a fundamental instrument as the clock. In the first place its length for *any* observer, as measured by the radar method, depends on the

² *Ibid.*, p. 39.

³ *Ibid.*, p. 41.

⁴ *Ibid.*, pp. 42–43.

⁵ *Ibid.*, p. 46.

clock used by the observer; in the second place, different observers assign different lengths to it even if their clocks are congruent, owing to the fact that the test of simultaneity is a conventional one. The clock, on the other hand, once graduated, gives epochs at itself which are independent of convention.

Once we have set up a clock, arbitrarily graduated, distances for the observer using this clock become definite. If a rod, moved from one position of rest relative to this observer to another position of rest relative to the same observer, possesses in the two positions the same length, as measured by this observer using his own clock, as graduated, then the rod is said to have undergone a rigid-body-displacement by this clock. In this way we see that once we have fixed on a clock, a rigid-body-displacement becomes definable. But until we have provided a clock, there is no way of saying what we mean by a rigid body under displacement.⁶

Now, if Milne is to make good his criticism of Einstein by erecting the space-time structure of special relativity on alternative epistemological foundations, he must provide us with inertial systems by means of the resources of his light-signal kinematics as well as with the measures of length and time on which the kinematics of special relativity is predicated. This means that he must be able to characterize inertial systems *within* the confines of his epistemological program as some kind of dense assemblage of equivalent particle observers filling space such that each particle observer is at rest relative to and synchronous with every other. We have already seen that he was wholly in error in charging Einstein with lack of awareness of the conventionality of spatial congruence as "defined" by the rigid rod. But that, much more fundamentally, he is mistaken in believing to have erected the kinematics of special relativity on an epistemologically more satisfactory base than Einstein did will now be made clear by reference to the following result pointed out by L. L. Whyte:⁷ Using only light signals and temporal succession without either a solid rigid rod or an isochronous material clock, it is *not* possible to construct ordinary measures of length and time. For "a

⁶ *Ibid.*, pp. 47-48.

⁷ L. L. Whyte: "Light Signal Kinematics," *British Journal for the Philosophy of Science*, Vol. IV (1953), pp. 160-61.

physicist using only light signals cannot discriminate inertial systems from these subjected to arbitrary 4-D similarity transformations.⁸ The system of 'resting' mass-points which can be so identified may be arbitrarily expanding and/or contracting relatively to a rod, and these superfluous transformations can only be eliminated by using a rod or a clock."⁹

The significance of the result stated by Whyte is twofold.

First: If Milne dispenses with material clocks and bases his chronometry only on the congruences yielded by his light-signal clocks, then he cannot obtain inertial systems without a rigid rod in the following sense. The rigid rod is *not* needed for the definition of *spatial congruence* within the system but *is* required to assure that the distance between one particular pair of points connected by it *at one time* t_0 is the same as *at some later time* t_1 . In other words, the rod is rigid *at a given place* by remaining congruent to itself (by convention) as time goes on. And in this way the rod assures the time-constancy of the distance between the two given points connected by it. This reliance on the rigid rod thus involves the use of the definition of simultaneity. Hence, if Milne were right in charging that the use of a rigid rod is beset by philosophic difficulties, then he indeed would be incurring these liabilities no less than Einstein does. However, let us suppose that:

Second, Milne does use a material clock to define the time-metric at a space point and thereby to particularize his clock graduations to the kind required for the elimination of the unwanted reference systems described by L. L. Whyte. This procedure is a far cry from his purely topological clock which "involves only the existence of an *ego*"¹ in contradistinction to the rigid scale's involvement of a definition of simultaneity. And, in that case, his measurement of the equality of space intervals by means of the equality of the corresponding round-trip time-intervals involves the following conventions: (a) the tacit use of a definition of simultaneity of non-coinciding events. For although

⁸ For a brief account of similarity transformations, and a further articulation of Whyte's point here, cf. H. Reichenbach: *PST*, *op. cit.*, pp. 172-73.

⁹ L. L. Whyte: "Light Signal Kinematics," *op. cit.*, p. 161.

¹ E. A. Milne: *Modern Cosmology and the Christian Idea of God*, *op. cit.*, p. 46.

a round-trip time on a given clock does not, of course, itself require such a simultaneity criterion, the measurement of a spatial distance in an inertial system by means of this time does: the distance yielded by the round-trip time on a clock at A is the distance r_2 between A and B *at the time t_2 on the A-clock when the light pulse from A arrives at B* on its round-trip ABA, (b) successive equal differences in the readings of a given local clock are stipulated to be measures of equal time intervals and thereby of equal space intervals, and (c) equal differences on separated clocks of identical constitution are decreed to be measures of equal time intervals and thereby of equal space intervals.

To what extent then, if any, does Milne have a case against Einstein? It would appear from our analysis that the only justifiable criticism is not at all epistemological but concerns an innocuous point of axiomatic economy: once you grant Milne a material clock, he does not require the rigid rod at all, whereas Einstein utilizes the spatial congruence definition based on the rigid rod *in addition* to all of the conventions needed by Milne. Thus, Milne's kinematics, as supplemented by the use of a material clock, is constructed on a slightly narrower base of conventions than is Einstein's.²

It will be recalled from Chapter Four that if measurements of spatial and temporal extension are to be made by means of solid rods and material clocks, allowance must be made computationally for thermal and other perturbations of these bodies so that they can define rigidity and isochronism. Calling attention to this fact and believing Milne's light-signal kinematics to be essentially successful, L. Page deemed Milne's construction more adequate than Einstein's, writing:

the original formulation of the relativity theory was based on undefined concepts of space and time intervals which could not be identified unambiguously with actual observations. Recently Milne has shown how to supply the desired criterion [of rigidity and isochronism] by erecting the space-time structure on the foundations of a constant light-signal velocity.³

² The preceding critique of Milne supplants my earlier brief critique ("The Philosophical Retention of Absolute Space in Einstein's General Theory of Relativity," *The Philosophical Review*, Vol. LXVI [1957], pp. 531-33) in which I *misinterpreted* Milne's arguments as indicative of lack of appreciation on his part of the conventionality of *temporal* congruence.

³ L. Page and N. I. Adams: *Electrodynamics*, *op. cit.*, pp. 78-79.

It is apparent in the light of our appraisal of Milne's kinematics that Page's claim is vitiated by Milne's need for a rigid rod or material clock as specified.

It should be noted, however, as Professor A. G. Walker has pointed out to me, that if Milne's construction is interpreted as applying *not* to special relativity kinematics but to his cosmological world model, then our criticisms are no longer pertinent. In terms of his logarithmically related τ and t scales of time, it turns out that upon measuring distances by the specified chronometric convention, the galaxies are at relative rest in τ -scale kinematics and in uniform relative motion in the t -scale. Each of these time scales is unique up to a trivial change of units, and their associated descriptions of the cosmological world are equivalent in Reichenbach's sense. In this *cosmological* context, the problem of eliminating the superfluous transformations mentioned by Whyte therefore does not arise.

Chapter 14

HAS THE GENERAL THEORY OF RELATIVITY REPUDIATED ABSOLUTE SPACE?

The literature of recent decades on the philosophy and history of science has nurtured and given wide currency to a myth concerning the present status of the dispute between the absolutistic and relativistic theories of space. In particular, that literature is rife with assertions that the post-Newtonian era has witnessed "the final elimination of the concept of absolute space from the conceptual scheme of modern physics"¹ by Einstein's general theory of relativity and that the Leibniz-Huyghens polemic against Newton and Clarke has thus been triumphantly vindicated.² In this vein, Philipp Frank recently reached the following verdict on Einstein's success in the implementation of Ernst Mach's program for a *relativistic* account of the *inertial* properties of matter: "Einstein started a new analysis of Newtonian mechanics which eventually vindicated Mach's reformulation [of Newtonian mechanics]."³

I shall now show that the history of the GTR does not at all bear out the widespread view set forth in the quotations from Max Jammer and Philipp Frank. And it will then become apparent in what precise sense there is ample justification for Einstein's own admission of 1953 as follows: the supplanting of the concept

¹ M. Jammer: *Concepts of Space*, *op. cit.*, p. 2.

² A very useful modern edition by H. G. Alexander of *The Leibniz-Clarke Correspondence* has been published in 1956 by the Manchester University Press and by the Philosophical Library in New York.

³ P. Frank: *Philosophy of Science* (Englewood Cliffs, New Jersey: Prentice-Hall; 1957), p. 153.

of absolute space is "a process which is probably by no means as yet completed."⁴

Mach had urged against Newton that both translational and rotational inertia are intrinsically dependent on the large-scale distribution and relative motion of matter. Assuming the indefinite extensibility of terrestrial axes to form an unlimited Euclidean rigid system S_* , the rotational motion of the stars seemed to be clearly defined with respect to S_* . Unfortunately, however, the GTR was not entitled to make use of S_* : the linear velocity of rotating mass points increases with the distance from the axis of rotation, and hence the existence of a system S_* of unrestricted size would allow *local* velocities greater than that of light, in contravention of the requirement of the local validity of the STR. But to deny, as the GTR therefore must, that S_* can extend even as far as the planet Neptune is to assert that the Machian concept of the *relative* motion of the earth and the stars is no more meaningful physically than the Newtonian bugaboo of the *absolute* rotation of a solitary earth in a space which is structured independently of any matter that it might contain accidentally and indifferently!⁵ Accordingly, the earth must be held to rotate *not* relative to the stars but with respect to the local "star-compass" formed at the earth by stellar light rays whose paths are determined by the local *metrical field*.

At Einstein's hands, Mach's thesis underwent not only this modification but also the following generalization: Einstein found that *both* the geometry of material rods and clocks *and* the inertial behavior of particles and light in the context of that geometry are functionally related to the same physical quantities. Probably unaware at the time that Riemann had previously conjectured the dependence of the geometry of physical space on the action of matter via a different line of reasoning,⁶ Einstein

⁴ A. Einstein: Foreword to Max Jammer, *Concepts of Space*, *op. cit.*, p. 15.

⁵ For details, see H. Weyl: "Massenträgheit und Kosmos," *Naturwissenschaften*, Vol. XII (1924), p. 197. See also F. A. E. Pirani: "On the Definition of Inertial Systems in General Relativity," in *Bern Jubilee of Relativity Theory*, suppl. IV of *Helvetica Physica Acta* (Basel: Birkhäuser Verlag; 1956), pp. 198–203.

⁶ Cf. B. Riemann: *Über die Hypothesen welche der Geometrie zu Grunde liegen*, edited by H. Weyl (3rd ed., Berlin: Julius Springer; 1923), pp. 3 and 20. The reader will find a brief account of the relevant part of Riemann's reasoning in Chapters One and Fifteen of the present book.

named his own organic fusion of Riemann's and Mach's ideas "Mach's Principle."⁷ And he sought to implement that principle by requiring that the metrical field given by the quantities g_{ik} be *exhaustively* determined by properties and relations of ponderable matter and energy specified by the quantities T_{ik} . On this conception a single test particle would have no inertia whatever if all other matter and energy were either annihilated or moved indefinitely far away.

But when the problem of solving the nonlinear partial differential equations which connect the derivatives of the g_{ik} to the T_{ik} was confronted, it became apparent that, far from having been exorcised by the GTR, the ghost of Newton's absolute space is nothing less than a haunting incubus. For to obtain a solution of these equations, it is necessary to supply the boundary conditions "at infinity." And to assume, as is done in Schwarzschild's solution, that there are certain preferred coordinate systems in which the g_{ik} have the Lorentz-Minkowski values at infinity is to violate Mach's Principle in the following twofold sense: first, the boundary conditions at infinity then assume the role of Newton's absolute space, since it is not the influence of matter that determines what coordinate systems at infinity are the Galilean ones of special relativity; and second, instead of being the *source* of the *total* structure of space-time, matter then merely *modifies* the latter's otherwise autonomously flat structure. In 1916 Einstein first attempted to avoid this most unwelcome consequence by reluctantly altering the above field equations through the introduction of the cosmological constant λ , which yielded a solution in which space was closed (finite). But this rather forced step did not provide an escape from the troublesome philosophical difficulties that had cropped up in the boundary conditions at infinity, since these difficulties reappeared when W. de Sitter showed that the now *modified* equations violated Mach's Principle by allowing a universe essentially devoid of *matter to have* a definitely structured space-time. The attempt to dispose of the difficulty at infinity by laying down the finitude of space as a *boundary condition* governing the solution of the *unmodified* field equations is unavailing for the purpose of rescu-

See correction in
Append. § 38

⁷ A. Einstein: "Prinzipielles zur allgemeinen Relativitätstheorie," *op. cit.*, p. 241.

ing Mach's Principle as it was originally conceived, since such a speculative assumption involves a nonintrinsic connection between the over-all structure of space and the properties of matter. In 1951 the Machian hope of subordinating space-time ontologically to matter was further dashed when A. H. Taub showed that there are conditions under which the *unmodified* field equations yield *curved* space in the absence of matter.⁸

These results inescapably raise the question of whether the failure of the GTR to implement Mach's Principle is to be regarded as an inadequacy on the part of that theory or as a basis for admitting that the GTR was right in philosophically retaining Newton's absolute space to a significant extent, thinly disguised under new structural trappings. Einstein's own attitude in his last years seems to have been one of unmourning abandonment of Mach's Principle. His reason appears to have been that although matter provides the *epistemological* basis for the metrical field, this fact must *not* be held to confer *ontological* primacy on matter over the field: matter is merely part of the field rather than its source.⁹

See Append. §40

This is indeed a very far cry from, nay the very antithesis of, Max Jammer's "final elimination of the concept of absolute space from the conceptual scheme of modern physics."¹ In fact, Jammer himself quotes a recent passage from Einstein in which Einstein says that if the *space-time field* were removed, there would be no space.² Yet Jammer gives no indication whatever that this is a drastically different thesis from Einstein's earlier one that if all

⁸ A. H. Taub: "Empty Space-Times Admitting a Three Parameter Group of Motions," *Annals of Mathematics*, Vol. LIII (1951), p. 472.

⁹ For a discussion of the status of Einstein's program of field theory, see J. Callaway: "Mach's Principle and Unified Field Theory," *Physical Review*, Vol. XCVI (1954), p. 778. For an alternative theory of gravitation inspired by the aim of strict conformity to Mach's Principle but incomplete in other respects, see D. W. Sciama: "On the Origin of Inertia," *Monthly Notices of the Royal Astronomical Society*, Vol. CXIII (1953), p. 35, and "Inertia," *Scientific American*, Vol. CXCVI (February, 1957), pp. 99-109. Cf. also F. A. Kaempffer: "On Possible Realizations of Mach's Program," *Canadian Journal of Physics*, Vol. XXXVI (1958), pp. 151-59, and O. Klein: "Mach's Principle and Cosmology in their Relation to General Relativity," in *Recent Developments in General Relativity* (Warsaw: Polish Scientific Publishers; 1962), pp. 293-302.

¹ M. Jammer: *Concepts of Space*, *op. cit.*, p. 2.

² *Ibid.*, p. 172.

matter were annihilated, then metric space would vanish as well.³

It is now clear that the GTR *cannot* be said to have resolved the controversy between the absolutistic and relativistic conceptions of space in favor of the latter on the issue of the implementation of Mach's Principle. Instead, the current state of knowledge supports the following summary assessment given in 1961 by the physicists C. Brans and R. H. Dicke:

The . . . view that the geometrical and inertial properties of space are meaningless for an empty space, that the physical properties of space have their origin in the matter contained therein, and that the only meaningful motion of a particle is motion relative to other matter in the universe has never found its complete expression in a physical theory. This picture is . . . old and can be traced from the writings of Bishop Berkeley⁴ to those of Ernst Mach.⁵ These ideas have found a limited expression in general relativity, but it must be admitted that, although in general relativity spatial geometries are affected by mass distributions, the geometry is not uniquely specified by the distribution. It has not yet been possible to specify boundary conditions on the field equations of general relativity which would bring the theory into accord with Mach's principle. Such boundary conditions would, among other things, eliminate all solutions without mass present.⁶

³ M. Jammer has since taken account of these criticisms on pp. 12 and 195 of the revised Harper Torchbook edition of his book, published in New York in 1960.

⁴ "G. Berkelsy: *The Principles of Human Knowledge*, Paragraphs 111-17, 1710-*De Motu* (1726)."

⁵ "E. Mach: *Conservation of Energy*, note No. 1, 1872 (reprinted by Open Court Publishing Company, LaSalle, Illinois, 1911), and *The Science of Mechanics*, 1883 (reprinted by Open Court Publishing Company, LaSalle, Illinois, 1902, Chapter II, Sec. VI)."

See Append. § 39

⁶ C. Brans and R. H. Dicke: "Mach's Principle and a Relativistic Theory of Gravitation," *The Physical Review*, Vol. CXXIV (1961), p. 925. See also R. H. Dicke: "Mach's Principle and Invariance Under Transformation of Units," *The Physical Review*, Vol. CXXV (1962), p. 2163, and "The Nature of Gravitation," in L. V. Berkner and H. Odishaw (eds.) *Science in Space* (New York: McGraw-Hill Book Company; 1961), Chapter iii, Sec. 3.1, "Mach's Principle," pp. 93-95. For an account of statements which might be regarded as modified versions of Mach's Principle and which are valid in the GTR, see C. H. Brans: "Mach's Principle and the Locally Measured

The difficulties encountered by the attempt to incorporate Mach's Principle as originally conceived into the GTR have most recently prompted two kinds of responses from leading investigators, which illustrate the lack of a uniform conception of this principle. Brans and Dicke⁷ have put forward a *modified* relativistic theory of gravitation which is apparently compatible with Mach's principle, and is closely related to the theory of P. Jordan.⁸ But J. A. Wheeler has articulated the important modifications which must be made in the original program of Mach's Principle, if Mach's ideas are to preserve their relevance to the GTR in its current state. Wheeler's substantial reformulation of Mach's Principle is as follows: "the specification of a sufficiently regular closed three-dimensional geometry at two immediately succeeding instants, and of the density and flow of mass-energy, is to determine the geometry of space-time, past, present, and future, and thereby the inertial properties of every infinitesimal test particle."⁹ On Wheeler's view then, Mach's Principle can be implemented in the GTR in the following drastically altered form: if we are *given* (1) that the *three*-dimensional geometry of space at some initial instant and at some closely succeeding instant does *not* extend to infinity and does *not* show infinite curvature, and (2) the distribution of mass and mass-flow, then the *four*-dimensional geometry of space-time or the "geometrodynamics" and hence the inertial properties of infinitesimal test particles are thereby determined. For Wheeler then, the modified form of Mach's Principle simply *requires ab initio* that the universe be *spatially* closed or finite. In this way, it constitutes a principle for *selecting* out of the many conceivable solutions of Einstein's field equations those for which the three-geometry at a given

Gravitational Constant in General Relativity," *The Physical Review*, Vol. CXXV (1962), p. 396.

⁷ C. Brans and R. H. Dicke: "Mach's Principle and a Relativistic Theory of Gravitation," *op. cit.*

⁸ P. Jordan: *Schwerkraft und Weltall* (Braunschweig: Friedrich Vieweg und Sohn; 1955).

⁹ J. A. Wheeler: "Mach's Principle as a Boundary Condition for Einstein's Field Equations and as a Central Part of the 'Plan' of General Relativity," a Report given at the Conference on Relativistic Theories of Gravitation, Warsaw, Poland, July, 1962.

instant is closed and free from singularity, thereby making possible the determination of the four-geometry and of the inertial behavior of infinitesimal test particles.¹

¹ J. A. Wheeler ("The Universe in the Light of General Relativity," *The Monist*, Vol. XLVII, No. 1 [1962], pp. 40-76) has given a very brief statement of the meaning of his reformulation of Mach's Principle as applied to a universe which is empty of all "real" mass in the sense of the vision of Clifford and Einstein. For details on the latter universe, see C. W. Misner and J. A. Wheeler: "Geometrodynamics," *Annals of Physics*, Vol. II (1957), pp. 525-614; J. A. Wheeler: "Curved Empty Space-Time As the Building Material of the Physical World: An Assessment," in: E. Nagel, P. Suppes, and A. Tarski (eds.) *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress* (Stanford: Stanford University Press; 1962), pp. 361-74, and J. G. Fletcher: "Geometrodynamics," in: L. Witten (ed.) *Gravitation* (New York: John Wiley & Sons, Inc.; 1962), Chapter xx, pp. 412-37.

Chapter 15

PHILOSOPHICAL CRITIQUE OF WHITEHEAD'S THEORY OF RELATIVITY

Einstein's theory of relativity was probably the most important influence on Whitehead's philosophy of science. But Whitehead's endeavor to reinterpret and modify Einstein's STR and GTR in terms of the categories of his own natural philosophy issued in two important philosophical divergences from Einstein: first, as will be recalled from Chapter Twelve, Whitehead erects the STR on his espousal of a *sensory absolute simultaneity* for any given inertial system in opposition to Einstein's theses on simultaneity, and second, Whitehead repudiates the GTR, because he rejects on epistemological grounds Riemann's conception of the relation between geometry and physics, which Einstein had attempted to weave into the logical fabric of the GTR via Mach's Principle, as explained in Chapter Fourteen.

In Chapter Twelve, I set forth my reasons for regarding Whitehead's rival to Einstein's conception of simultaneity as utterly untenable.¹ It remains therefore in the present chapter to give a critical appraisal of Whitehead's grounds for renouncing Riemann's view of the relation between geometry and physics, a view which permeates Einstein's GTR. Says Whitehead:

I deduce that our experience requires and exhibits a basis of uniformity. . . . This conclusion entirely cuts away the causal

¹For a statement of my objections to R. M. Palter's recent defense (*Whitehead's Philosophy of Science, op. cit.*, pp. 34-41) of Whitehead's doctrine of simultaneity, cf. A. Grünbaum: "Whitehead's Philosophy of Science," *The Philosophical Review*, Vol. LXXI (1962), pp. 222-24.

heterogeneity [of the geometry] . . . which is the essential of Einstein's later theory [of *general* relativity].

It is inherent in my theory to maintain the old division between physics and geometry. Physics is the science of the contingent relations of nature and geometry expresses its uniform relatedness.²

Thus, in opposition to the GTR's spatially and temporally *variable* geometry, Whitehead affirms the *uniformity* of the world's geometry, thereby claiming that the geometry is either Euclidean or one of the non-Euclidean geometries of *constant* negative or positive curvature.

To discern the philosophical basis for Whitehead's position, we must be mindful of the reasoning which issues in Riemann's conception of the relation between geometry and physics. Riemann had drawn the revolutionary conclusion that the metric geometry prevailing in physical space with respect to a specified congruence standard is acquired by that space, as Weyl puts it, "*only through the advent of the material content filling it and determining its metric relations.*"³ In other words, "Riemann rejects the opinion that had prevailed up to his own time, namely, that the metrical structure of space is fixed and inherently independent of the physical phenomena for which it serves as a background, and that the real content takes possession of it as of residential flats."⁴ And, as the reader will recall from the statement by Riemann quoted in Chapter One, Section B, Riemann reaches this conclusion *on the basis of his conception of the status of spatial congruence* as follows. In the continuous manifold of space (or of time), the measure of an interval must be provided by a congruence standard which "must come from somewhere else [i.e., from outside the manifold itself]"⁵ by convention. *But the coincidence behavior of that standard under transport is left indeterminate by the structure of the spatial (or temporal) manifold itself.* Hence Riemann infers that if "the actual things forming the groundwork of a space" form a continuous manifold, then

² A. N. Whitehead: *The Principle of Relativity*, *op. cit.*, pp. 5 and 6.

³ H. Weyl: *Space-Time-Matter*, *op. cit.*, p. 98.

⁴ *Ibid.*

⁵ B. Riemann: *On The Hypotheses Which Lie At The Foundations of Geometry*, *op. cit.*, pp. 424-25.

"the basis of metric relations must be sought for outside that actuality, in colligating forces that operate upon it."⁶

It will now become apparent that it was Whitehead's anti-bifurcationist and perceptualistic conception of spatial and temporal congruence—which we had reason to *reject* in Chapter One, Section G—that is the *fons* and *origo* of his renunciation of Riemann's reasoning and hence of Einstein's countenancing of a *variable* geometry in the GTR. For Whitehead's elaboration of the geometry and kinematics ingredient in his rival theory of relativity via his method of extensive abstraction⁷ rests on the contention that the unique spatial and temporal congruences provided by our *sensory* contents are such as to require the spatial and temporal *uniformity* of the world's geometry. But if I succeeded in demonstrating the untenability of Whitehead's theses regarding congruence in Chapter One, Section G, then his argument for the uniformity of the world's geometry is clearly gratuitous.

Palter has pointed out⁸ that Whitehead adduces his account of sense perception in the mode of pure relatedness as an argument for the uniformity of space, making no mention of congruence in that context. And thus it might seem that I have erred in representing Whitehead as having rested his case for uniformity on the perception of congruence. But there is a compelling mathematical reason for rejecting Palter's view that Whitehead intended his avowal of uniformity to be independent of his account of congruence and that he regarded the latter as relevant only to the determination of which one of the three kinds of uniform geometries obtains in nature. As a mathematician, Whitehead must have been aware that an assertion of uniform metrical relatedness has no meaning without at least tacit reference to a particular criterion of congruence: if space were to have a variable geometry with respect to a certain

⁶ *Ibid.*, p. 425.

⁷ My reasons for believing that Whitehead's method of extensive abstraction is a failure within the framework of his perceptualistic epistemology are given in A. Grünbaum: "Whitehead's Method of Extensive Abstraction," *The British Journal for the Philosophy of Science*, Vol. IV (1953), pp. 215–26.

⁸ R. M. Palter: *Whitehead's Philosophy of Science*, *op. cit.*, pp. 25–27 and 32–33.

congruence, then there would always be infinitely many other congruences that would impart a uniform geometry to it, and conversely. Thus, for example, although the surface of an egg exhibits a variable geometry (Gaussian curvature) with respect to the standard metrization, there are infinitely many other congruences with respect to which that very surface would exhibit the uniform geometry imparted to the surface of a sphere by the standard metrization. Hence, without a specification of the congruence criterion, the mere claim of spatial uniformity places no restrictions on the coincidence behavior of transported rods and does not rule out the variable geometry of Einstein's GTR. It seems unavoidable, therefore, to conclude that Whitehead intended his arguments for uniform relatedness to stand or fall with his theses regarding congruence.

Since I have argued that these latter theses are quite unsound, I must reach the following verdict in regard to Whitehead's rival to Einstein's GTR: if it should turn out that the uniformity of space affirmed by his alternative to Einstein's GTR is strongly supported by future empirical findings, this will *not* be because of the *philosophical* reasons put forward by Whitehead.

Bibliography

- Aharoni, J.: *The Special Theory of Relativity*. Oxford: Oxford University Press; 1959.
- Akhiezer, N. I.: *The Calculus of Variations*. New York: Blaisdell Publishing Company; 1962.
- Alexander, H. G.: *The Leibniz-Clarke Correspondence*. Manchester: Manchester University Press; 1956, and New York: Philosophical Library; 1956.
- Alexander, H. L.: *Reactions with Drug Therapy*. Philadelphia: W. B. Saunders Company; 1955.
- Aristotle: *On Generation and Corruption*. Book I, Ch. ii, 316^a15–317^a17.
- Baker, B. and E. T. Copson: *The Mathematical Theory of Huyghens's Principle*. Oxford: Oxford University Press; 1939.
- Baldus, R.: *Nichteuklidische Geometrie*. Edited by F. Löbell. Third revised edition. Berlin: Walter de Gruyter and Company; 1953. Sammlung Göschel, Vol. CMLXX.
- Barankin, E. W.: "Heat Flow and Non-Euclidean Geometry." *American Mathematical Monthly*, Vol. XLIX (1942).
- Baruch, J. J.: "Horological Accuracy: Its Limits and Implications." *American Scientist*, Vol. XLVI (1958).
- Baum, W. A.: "Photoelectric Test of World Models." *Science*, Vol. CXXXIV (1961).
- Bell, E. T.: *The Development of Mathematics*. New York: McGraw-Hill Book Company; 1945.
- Bergmann, G.: "Review of M. Born's *Natural Philosophy of Cause and Chance*." *Philosophy of Science*, Vol. XVII (1950).
- Bergmann, H.: *Der Kampf um das Kausalgesetz in der jüngsten Physik*. Braunschweig: F. Vieweg and Son; 1929.
- Bergmann, P.: *Introduction to the Theory of Relativity*. New York: Prentice-Hall; 1946.
- : "Fifty Years of Relativity." *Science*, Vol. CXXIII (1956).
- Bergson, H.: *Creative Evolution*. New York: Random House, Inc.; 1944.
- : *Matière et Mémoire*. Geneva: A. Skira; 1946.
- Bergstrand, E.: "Determination of the Velocity of Light." *Handbuch der Physik*. Edited by W. Flügge. Berlin: J. Springer; 1956. Vol. XXIV.

- Berkeley, G.: *The Principles of Human Knowledge*. Paragraphs 111-17, 1710-*De Motu* (1726).
- Birkhoff, G. D.: *The Origin, Nature and Influence of Relativity*. New York: The Macmillan Company; 1925.
- Black, M.: "The 'Direction' of Time." *Analysis*, Vol. XIX (1959).
- : "Review of G. J. Whitrow's *The Natural Philosophy of Time*." *Scientific American*, Vol. CCVI (1962).
- Blank, A. A.: "Analysis of Experiments in Binocular Space Perception." *Journal of the Optical Society of America*, Vol. XLVIII (1958).
- : "Axiomatics of Binocular Vision. The Foundations of Metric Geometry in Relation to Space Perception." *Ibid.*
- : "The Geometry of Vision." *The British Journal of Physiological Optics*, Vol. XIV (1957).
- : "The Luneburg Theory of Binocular Perception." *Psychology, A Study of a Science*. Edited by S. Koch. New York: McGraw-Hill Book Company, Inc.; 1958. Study I, Vol. I.
- : "The Luneburg Theory of Binocular Visual Space." *Journal of the Optical Society of America*, Vol. XLIII (1953).
- : "The non-Euclidean Geometry of Binocular Visual Space." *Bulletin of the American Mathematical Society*, Vol. LX (1954).
- Blatt, J. M.: "Time Reversal." *Scientific American* (August 1956).
- Blum, H. F.: *Time's Arrow and Evolution*. Second edition. Princeton: Princeton University Press; 1955.
- Bohm, D.: *Quantum Theory*. New York: Prentice-Hall; 1951.
- Bolza, O.: *Lectures on the Calculus of Variations*. New York: G. E. Stechert; 1946.
- Bolzano, B.: *Paradoxes of the Infinite*. Edited by D. A. Steele. New Haven: Yale University Press; 1951.
- Bonch-Bruевич, A. M.: "On the Direct Experimental Verification of the Second Postulate of the Special Relativity Theory." *Physics Express* (November 1960), Vol. III.
- Bondi, H.: *Cosmology*. Second Edition. Cambridge: Cambridge University Press; 1961.
- : "Relativity and Indeterminacy." *Nature*, Vol. CLXIX (1952).
- Bonola, R.: *Non-Euclidean Geometry*. New York: Dover Publications, Inc.; 1955.
- Born, M.: "Physical Reality." *Philosophical Quarterly*, Vol. III (1953).
- : *Natural Philosophy of Cause and Chance*. Oxford: Oxford University Press; 1949.
- Boyer, C. B.: *The Concepts of the Calculus*. New York: Hafner Publishing Company; 1949.
- Braithwaite, R. B.: "Axiomatizing a Scientific System by Axioms in the Form of Identification." *The Axiomatic Method*. Edited by L. Henkin, P. Suppes, and A. Tarski. Amsterdam: North Holland Publishing Company; 1959.
- Brans, C. H. and Dicke, R. H.: "Mach's Principle and a Relativistic Theory of Gravitation." *The Physical Review*, Vol. CXXIV (1961).

- Brans, C. H.: "Mach's Principle and the Locally Measured Gravitational Constant in General Relativity." *Ibid.*, Vol. CXXV (1962).
- Bridgman, P. W.: "The Nature of Physical 'Knowledge.'" *The Nature of Physical Knowledge*. Edited by L. W. Friedrich. Bloomington: Indiana University Press; 1960.
- : *Reflections of a Physicist*. New York: Philosophical Library; 1950.
- : "Reflections on Thermodynamics." *American Scientist*, Vol. XLI (1953).
- : "Some Implications of Recent Points of View in Physics." *Revue Internationale de Philosophie*, Vol. III, No. 10 (1949).
- : *A Sophisticate's Primer of the Theory of Relativity*. Middletown, Conn.: Wesleyan University Press; 1962.
- Brillouin, L.: *Science and Information Theory*. New York: Academic Press; 1956.
- Brouwer, C.: "The Accurate Measurement of Time." *Physics Today*, Vol. IV (1951).
- Brown, F. A., Jr.: "Biological Clocks and the Fiddler Crab." *Scientific American*, Vol. CXC (1954).
- : "Living Clocks." *Science*, Vol. CXXX (1959).
- : "Response to Pervasive Geophysical Factors and the Biological Clock Problem." *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. XXV (1960).
- : "The Rhythmic Nature of Animals and Plants." *American Scientist*, Vol. XLVII (1959).
- Burington, R. S. and Torrance, C. C.: *Higher Mathematics*. New York: McGraw-Hill Book Company, Inc.; 1939.
- Callaway, J.: "Mach's Principle and Unified Field Theory." *Physical Review*, Vol. XCVI (1954).
- Cantor, G.: *Gesammelte Abhandlungen*. Edited by E. Zermelo. Berlin: J. Springer; 1932.
- Capek, M.: *The Philosophical Impact of Contemporary Physics*. Princeton: D. Van Nostrand Company, Inc.; 1961.
- Carathéodory, C.: "Untersuchungen über die Grundlagen der Thermodynamik." *Mathematische Annalen* (1909).
- Carnap, R.: *Abriss der Logistik*. Vienna: J. Springer; 1929.
- : "Adolf Grünbaum on The Philosophy of Space and Time." *The Philosophy of Rudolf Carnap*. Edited by P. A. Schilpp. LaSalle: Open Court Publishing Company; 1963.
- : *Der Raum*. Berlin: Reuther und Reichard; 1922.
- : *Introduction to Symbolic Logic and Its Applications*. New York: Dover Publications, Inc.; 1958.
- : *Symbolische Logik*. Vienna: J. Springer; 1954.
- : "Über die Abhängigkeit der Eigenschaften des Raumes von denen der Zeit." *Kantstudien*, Vol. XXX (1925).
- Cassirer, E.: *Zur Einsteinschen Relativitätstheorie*. Berlin: Bruno Cassirer Verlag; 1921.

- Cedarholm, J. P. and Townes, C. H.: "A New Experimental Test of Special Relativity." *Nature*, Vol. CLXXXIV, No. 4696 (1959).
- Chandrasekhar, S. and Wright, J. P.: "The Geodesics in Godel's Universe." *Proceedings of the National Academy of Sciences*, Vol. XLVIII (1961).
- Cherry, E. C.: "The Communication of Information." *American Scientist*, Vol. XL (1952).
- Clemence, G. M.: "Astronomical Time." *Review of Modern Physics*, Vol. XXIX (1957).
- : "Dynamics of the Solar System." *Handbook of Physics*. Edited by E. Condon and H. Odishaw. New York: McGraw-Hill Book Company, Inc.; 1958.
- : "Ephemeris Time." *Astronomical Journal*, Vol. LXIV (1959) and *Transactions of the International Astronomical Union*, Vol. X (1958).
- : "Time and Its Measurement." *American Scientist*, Vol. XL (1952).
- Clifford, W. K.: *The Common Sense of the Exact Sciences*. New York: Dover Publications, Inc.; 1955.
- Courant, R.: *Vorlesungen über Differential und Integralrechnung*. Berlin: J. Springer; 1927.
- Courant, R. and Robbins, H.: *What is Mathematics?* New York: Oxford University Press; 1941.
- Cramér, H.: *Mathematical Methods of Statistics*. Princeton: Princeton University Press; 1946.
- d'Abro, A.: *The Evolution of Scientific Thought from Newton to Einstein*. New York: Dover Publications, Inc.; 1950.
- Davies, R. O.: "Irreversible Changes: New Thermodynamics From Old." *Science News*, No. 28 (May 1953).
- de Beauregard, O. C.: "Complémentarité et Relativité," *Revue Philosophique*, Vol. CXLV (1955).
- : "L'Irréversibilité Quantique, Phénomène Macroscopique." *Louis de Broglie, Physicien et Penseur*. Paris: Albin Michel; 1953.
- : *Théorie Synthétique de la Relativité Restreinte et des Quanta*. Paris: Gauthier-Villars; 1957.
- Dedekind, R.: *Essays on the Theory of Number*. Chicago: Open Court Publishing Company; 1901.
- Dicke, R. H.: "Mach's Principle and Invariance under Transformation of Units." *Physical Review*, Vol. CXXV (1962).
- : "The Nature of Gravitation." *Science in Space*. Edited by L. V. Berkner and H. Odishaw. New York: McGraw-Hill Book Company; 1961.
- Dingle, H.: "Falsifiability of the Lorentz-Fitzgerald Contraction Hypothesis." *The British Journal for the Philosophy of Science*, Vol. X (1959).
- : *The Sources of Eddington's Philosophy*. London: Cambridge University Press; 1954.
- Dingler, H.: "Die Rolle der Konvention in der Physik." *Physikalische Zeitschrift*, Vol. XXIII (1922).
- Driesch, H.: *Philosophische Gegenwartsfragen*. Leipzig: E. Reinicke; 1933.
- du Bois-Reymond, P.: *Die Allgemeine Funktionentheorie*. Tübingen: Lauppische Buchhandlung; 1882. Vol. I.

- Duhem, P.: *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press; 1954.
- Eddington, A. S.: *The Nature of the Physical World*. New York: The Macmillan Company; 1928.
- : *Space, Time and Gravitation*. Cambridge: Cambridge University Press; 1953.
- Edel, A.: *Aristotle's Theory of the Infinite*. New York: Columbia University Press; 1934.
- Ehrenfest, P.: "In What Way Does It Become Manifest In The Fundamental Laws of Physics That Space Has Three Dimensions?" *Proceedings of the Amsterdam Academy*, Vol. XX (1917). (German translation: "Welche Rolle spielt die Dreidimensionalität des Raumes in den Grundgesetzen der Physik." *Annalen der Physik*, Vol. LXI [1920].)
- Ehrenfest, P. and T.: "Begriffliche Grundlagen der statistischen Auffassung in der Mechanik." *Encyklopädie der mathematischen Wissenschaften*, IV, 2, II.
- Einstein, A.: "Autobiographical Notes." *Albert Einstein: Philosopher-Scientist*. Edited by P. A. Schilpp. New York: Tudor Publishing Company; 1949.
- : "Die Relativitätstheorie." *Physik*. Edited by E. Warburg. Leipzig: Teubner; 1915.
- : Foreword to Max Jammer, *Concepts of Space*. Cambridge: Harvard University Press; 1954.
- : "The Foundations of the General Theory of Relativity." *The Principle of Relativity, A Collection of Original Memoirs*. New York: Dover Publications, Inc.; 1952.
- : *Geometrie und Erfahrung*. Berlin: J. Springer; 1921.
- : *The Meaning of Relativity*. Princeton: Princeton University Press; 1955.
- : "On the Electrodynamics of Moving Bodies." *The Principle of Relativity, A Collection of Original Memoirs*. New York: Dover Publications, Inc.; 1952.
- : "On the Influence of Gravitation on the Propagation of Light." *Ibid.*
- : "Prinzipielles zur allgemeinen Relativitätstheorie." *Annalen der Physik*, Vol. LV (1918).
- : "Reply to Criticisms." *Albert Einstein: Philosopher-Scientist*. Edited by P. A. Schilpp. Evanston: The Library of Living Philosophers; 1949.
- : "Über die Entwicklung unserer Anschauungen über die Konstitution und das Wesen der Strahlung." *Physikalische Zeitschrift*, Vol. X (1910).
- Eisenhart, L. P.: *Coordinate Geometry*. New York: Dover Publications, Inc.; 1960.
- : *An Introduction to Differential Geometry*. Princeton: Princeton University Press; 1947.
- : *Riemannian Geometry. Ibid.*, 1949.

- Epstein, P.: "The Time Concept in Restricted Relativity," and "A Rejoinder." *American Journal of Physics*, Vol. X (1942).
- : "Critical Appreciation of Gibbs' Statistical Mechanics." *A Commentary on the Scientific Writings of J. Willard Gibbs*. Edited by A. Haas. New Haven: Yale University Press; 1936. Vol. II.
- Essen, L. and Parry, J. V. L.: "An Atomic Standard of Frequency and Time Interval." *Nature*, Vol. CLXXVI (1955).
- Euclid: *The Thirteen Books of Euclid's Elements*. Trans. by T. L. Heath. Cambridge: Cambridge University Press; 1926.
- Feyerabend, P. K.: "Comments on Grünbaum's 'Law and Convention in Physical Theory.'" *Current Issues in the Philosophy of Science*. Edited by H. Feigl and G. Maxwell. New York: Holt, Rinehart and Winston; 1961.
- Fletcher, J. G.: "Geometrodynamics." *Gravitation*. Edited by L. Witten. New York: John Wiley & Sons, Inc.; 1962.
- Frank, P.: *Philosophy of Science*. Englewood Cliffs: Prentice-Hall; 1957.
- Fraenkel, A. A.: *Einleitung in die Mengenlehre*. New York: Dover Publications, Inc.; 1946.
- Fraenkel, A. A. and Bar-Hillel, Y.: *Foundations of Set Theory*. Amsterdam: North Holland Publishing Company; 1958.
- Freudenthal, H.: *Mathematical Reviews*, Vol. XXII (1961).
- : "Zur Geschichte der Grundlagen der Geometrie." *Nieuw Archief voor Wiskunde*, Vol. V (1957).
- Fürth, R.: "Prinzipien der Statistik." *Handbuch der Physik*. Edited by H. Geiger and K. Scheel. Berlin: J. Springer; 1929. Vol. IV.
- Gerhardt, K.: "Nichteuklidische Anschauung und optische Täuschungen." *Naturwissenschaften*, Vol. XXIV (1936).
- : "Nichteuklidische Kinematographie." *Naturwissenschaften*, Vol. XX (1932).
- Gerlach, W.: *Handbuch der Physik*. Berlin: Springer-Verlag; 1926.
- Gödel, K.: "An Example of a New Type of Cosmological Solutions of Einstein's Equations of Gravitation." *Reviews of Modern Physics*, Vol. XXI (1949).
- : "A Remark About the Relationship of Relativity Theory and Idealistic Philosophy." *Albert Einstein: Philosopher-Scientist*. Edited by P. A. Schilpp. Evanston: The Library of Living Philosophers; 1949.
- Gold, T.: "The Arrow of Time." *La Structure et L'Évolution de l'Univers*. Brussels: R. Stoops; 1958.
- Goodhard, C. B.: "Biological Time." *Discovery* (December 1957).
- Goodman, N.: *Fact, Fiction and Forecast*. Cambridge: Harvard University Press; 1955.
- Grünbaum, A.: "A Consistent Conception of the Extended Linear Continuum as an Aggregate of Unextended Elements." *Philosophy of Science*, Vol. XIX (1952).
- : *American Journal of Physics*, Vol. XXIV (1956).
- : "Causality and the Science of Human Behavior." *American Scientist*, Vol. XL (1952). (Reprinted in *Readings in the Philosophy of*

- Science. Edited by H. Feigl and M. Brodbeck. New York: Appleton-Century-Crofts, Inc.; 1953.)
- : "The Clock Paradox in the Special Theory of Relativity." *Philosophy of Science*, Vol. XXI (1954), and Vol. XXII (1955).
- : "Complementarity in Quantum Physics and Its Philosophical Generalization." *The Journal of Philosophy*, Vol. LIV (1957).
- : "Das Zeitproblem." *Archiv für Philosophie*, Vol. VII (1957).
- : "Messrs. Black and Taylor on Temporal Paradoxes." *Analysis*, Vol. XII (1952).
- : "Modern Science and the Refutation of the Paradoxes of Zeno." *The Scientific Monthly*, Vol. LXXXI (1955).
- : "Operationism and Relativity." *The Validation of Scientific Theories*. Boston: Beacon Press; 1957.
- : *Philosophy of Science*, Vol. XXII (1955).
- : "The Philosophical Retention of Absolute Space in Einstein's General Theory of Relativity." *The Philosophical Review*, Vol. LXVI (1957).
- : "Professor Dingle on Falsifiability: A Second Rejoinder." *British Journal for the Philosophy of Science*, Vol. XII (1961).
- : "Rejoinder to Feyerabend." *Current Issues in the Philosophy of Science*. Edited by H. Feigl and G. Maxwell. New York: Holt, Rinehart and Winston; 1961.
- : "Relativity and the Atomicity of Becoming." *The Review of Metaphysics*, Vol. IV (1950).
- : "Science and Ideology." *The Scientific Monthly*, Vol. LXXIX (1954).
- : "Science and Man." *Perspectives in Biology and Medicine*, Vol. V (1962).
- : *Scientific American*, Vol. CLXXXIX (December 1953).
- : "Some Highlights of Modern Cosmology and Cosmogony." *The Review of Metaphysics*, Vol. V (1952).
- : "Whitehead's Method of Extensive Abstraction." *The British Journal for the Philosophy of Science*, Vol. IV (1953).
- : "Whitehead's Philosophy of Science." *The Philosophical Review*, Vol. LXXI (1962).
- Hadamard, J.: *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. New Haven: Yale University Press; 1923.
- Halmos, P. R.: *Measure Theory*. New York: D. Van Nostrand Company, Inc.; 1950.
- Hanson, N. R.: "On the Symmetry Between Explanation and Prediction." *The Philosophical Review*, Vol. LXVIII (1959).
- Hardy, G. H.: *A Course of Pure Mathematics*. Ninth Edition. New York: The Macmillan Company; 1945.
- Hardy, L. H., Rand, G., Rittler, M. C., Blank, A. A., and Boeder, P.: *The Geometry of Binocular Space Perception*. New York: Columbia University College of Physicians and Surgeons; 1953.

- Hartshorne, C. and Weiss, P. (eds.): *The Collected Papers of Charles Sanders Peirce*. Cambridge: Harvard University Press; 1935. Vol. VI.
- Hasse, H. and Scholz, H.: *Die Grundlagenkrise der griechischen Mathematik*. Charlottenburg: Pan-Verlag; 1928.
- Heath, T. L.: *Mathematics in Aristotle*. Oxford: Oxford University Press; 1949.
- Hempel, C. G.: "Deductive Nomological vs. Statistical Explanation" *Minnesota Studies in the Philosophy of Science*, Vol. III. Minneapolis: University of Minnesota Press; 1962.
- Hempel, C. G. and Oppenheim, P.: "Studies in the Logic of Explanation." *Philosophy of Science*, Vol. XV (1948).
- Hertz, H.: *The Principles of Mechanics*. New York: Dover Publications, Inc.; 1956.
- Hilbert, D. and Bernays, P.: *Grundlagen der Mathematik*. Berlin: J. Springer; 1934.
- Hill, E. L. and Grünbaum, A.: "Irreversible Processes in Physical Theory." *Nature*, Vol. CLXXIX (1957).
- Hille, E.: *Functional Analysis and Semi-Groups*. New York: Mathematical Society Publications; 1948. Vol. XXXI.
- Hjelmslev, J.: "Die natürliche Geometrie." *Abhandlungen aus dem mathematischen Seminar der Hamburger Universität*, Vol. II (1923).
- Hoagland, H.: "Chemical Pacemakers and Physiological Rhythms." *Colloid Chemistry*. Edited by J. Alexander. New York. Vol. V (1944).
- : "The Physiological Control of Judgments of Duration: Evidence for a Chemical Clock." *The Journal of General Psychology*, Vol. IX (1933).
- Hobson, E. W.: *The Theory of Functions of a Real Variable*. New York: Dover Publications, Inc.; 1957. Vol. I.
- Hölder, O.: *Die Mathematische Methode*. Berlin: J. Springer; 1924.
- Holton, G.: "On the Origins of the Special Theory of Relativity." *American Journal of Physics*, Vol. XXVIII (1960).
- Hood, P.: *How Time is Measured*. London: Oxford University Press; 1955.
- Huntington, E. V.: *The Continuum and Other Types of Serial Order*. Second Edition. Cambridge: Harvard University Press; 1942.
- : "Inter-Relations Among the Four Principal Types of Order." *Transactions of the American Mathematical Society*, Vol. XXXVIII (1935).
- Huntington, E. V. and K. E. Rosinger: "Postulates for Separation of Point-Pairs (Reversible Order on a Closed Line)." *Proceedings of the American Academy of Arts and Sciences*. Boston. Vol. LXII (1932).
- Hutten, E. H.: *The Language of Modern Physics*. London: George Allen and Unwin, Ltd., and New York: The Macmillan Company; 1956.
- Jaffe, B.: *Michelson and the Speed of Light*. New York: Doubleday and Company; 1960.
- James, W.: *The Principles of Psychology*. New York: Dover Publications, Inc.; 1950.
- Jammer, M.: *Concepts of Space*. Cambridge: Harvard University Press; 1954.
- Jeffreys, H.: *The Earth*. Cambridge: Cambridge University Press; 1952.

- John, F.: "Numerical Solution of the Equation of Heat Conduction for Preceding Times." *Annali di Matematica Pura ed Applicata*, Vol. XL (1955).
- Jordan, P.: *Schwerkraft und Weltall*. Braunschweig: F. Vieweg und Sohn; 1955.
- Kaempffer, F. A.: "On Possible Realizations of Mach's Program." *Canadian Journal of Physics*, Vol. XXXVI (1958).
- Kennedy, R. J. and Thorndike, E. M.: *Physical Review*, Vol. XLII (1932).
- Klein, F.: *Elementary Mathematics From An Advanced Standpoint*. New York: Dover Publications, Inc.; 1939. Vol. II.
- : *Vorlesungen über Nicht-Euklidische Geometrie*. Berlin: Springer-Verlag; 1928.
- Klein, O.: "Mach's Principle and Cosmology in their Relation to General Relativity." *Recent Developments in General Relativity*. Warsaw: Polish Scientific Publishers; 1962.
- Landau, L. and Lifschitz, E.: *Statistical Physics*. Second Edition. Reading: Addison-Wesley; 1958.
- Landé, A.: "Axiomatische Begründung der Thermodynamik durch Carathéodory." *Handbuch der Physik*, Vol. IX (1926).
- : "The Logic of Quanta." *British Journal for the Philosophy of Science*, Vol. VI (1956).
- : "Optik und Thermodynamik." *Handbuch der Physik*. Berlin: J. Springer; 1928. Vol. XX.
- : "Wellenmechanik und Irreversibilität." *Physikalische Blätter*, Vol. XIII (1957).
- Lechalas, G.: *Étude sur l'espace et le temps*. Paris: Alcan Publishing Company; 1896.
- : "L'Axiome de libre Mobilité." *Revue de Métaphysique et de Morale*, Vol. VI (1898).
- Leclercq, R.: *Guide Théorique et Pratique de la Recherche Expérimentale*. Paris: Gauthier-Villars; 1958.
- Lee, H. D. P.: *Zeno of Elea*. Cambridge: Luzac and Company; 1936.
- Lefschetz, S.: *Introduction to Topology*. Princeton: Princeton University Press; 1949.
- Leibniz, G. W.: "Initia rerum mathematicorum metaphysica." *Mathematische Schriften*. Edited by C. J. Gerhardt. Berlin: Schmidt's Verlag; 1863.
- Lewin, K.: "Die zeitliche Geneseordnung." *Zeitschrift für Physik*, Vol. XIII (1923).
- Lewis, C. I.: *Analysis of Knowledge and Valuation*. LaSalle: Open Court Publishing Company; 1946.
- , and Langford, C. H.: *Symbolic Logic*. New York: The Century Company; 1932.
- Lorentz, H.: *The Theory of Electrons*. New York: Columbia University Press; 1909.
- Loschmidt, J.: "Über das Wärmegleichgewicht eines Systems von Körpern mit Rücksicht auf die Schwere." *Sitzungsberichte der Akademie der Wissenschaften*. Vienna, Vol. LXXIII (1876) and Vol. LXXV (1877).

- Ludwig, G.: "Der Messprozess." *Zeitschrift für Physik*, Vol. CXXXV (1953).
- : *Die Grundlagen der Quantenmechanik*. Berlin: J. Springer; 1954.
- : "Die Stellung des Subjekts in der Quantentheorie." *Veritas, Justitia, Libertas*. Festschrift zur 200-Jahr Feier der Columbia University. Berlin: Colloquium Verlag; 1954.
- : "Questions of Irreversibility and Ergodicity." *Ergodic Theories*. Edited by P. Caldivola. New York: Academic Press; 1961.
- : "Zum Ergodensatz und zum Begriff der makroskopischen Observablen. I." *Zeitschrift für Physik*, Vol. CL (1958).
- Lunenburg, R. K.: *Mathematical Analysis of Binocular Vision*. Princeton: Princeton University Press; 1947.
- : "Metric Methods in Binocular Visual Perception." *Studies and Essays*, Courant Anniversary Volume. New York: Interscience Publishers, Inc.; 1948.
- Luria, S.: "Die Infinitesimaltheorie der antiken Atomisten." *Quellen und Studien zur Geschichte der Mathematik, Astronomie, und Physik*. Berlin, 1933. Abteilung B: Studien, II.
- Lyons, H.: "Atomic Clocks." *Scientific American*, Vol. CXCVI (1957).
- Mach, E.: *Conservation of Energy*, 1872. Reprinted in LaSalle: Open Court Publishing Company; 1911.
- : *The Science of Mechanics*. 1883. Reprinted in LaSalle: Open Court Publishing Company; 1902.
- Margenau, H.: "Can Time Flow Backwards?" *Philosophy of Science*, Vol. XXI (1954).
- : "Measurements and Quantum States." *Philosophy of Science*, Vol. XXX (1963).
- Margenau, H. and Murphy, G. M.: *The Mathematics of Physics and Chemistry*. New York: D. Van Nostrand Company; 1943.
- Maritain, J.: *The Degrees of Knowledge*. London: G. Bles Company; 1937.
- Maxwell, J. C.: *Theory of Heat*. Sixth edition. New York: Longman's, Green, and Company; 1880.
- Mayr, E.: "Cause and Effect in Biology." *Science*, Vol. CXXXIV (1961).
- Mehlberg, H.: "Essai sur la théorie causale du temps." *Studia Philosophica*, Vol. I (1935), and Vol. II (1937).
- : "Physical Laws and Time's Arrow." *Current Issues in the Philosophy of Science*. Edited by H. Feigl and G. Maxwell. New York: Holt, Rinehart, and Winston; 1961.
- Menger, K.: *Dimensionstheorie*. Leipzig: B. G. Teubner; 1928.
- Michelson, A. A. and Gale, H. G.: "The Effect of the Earth's Rotation on the Velocity of Light." *Astrophysical Journal*, Vol. LXI (1925).
- Milham, W. I.: *Time and Timekeepers*. New York: The Macmillan Company; 1929.
- Milne, E. A.: *Kinematic Relativity*. Oxford: Oxford University Press; 1948.
- : *Modern Cosmology and the Christian Idea of God*. Oxford: Clarendon Press; 1952.
- : *Sir James Jeans*. Cambridge: Cambridge University Press; 1952.

- Misner, C. W. and Wheeler, J. A.: "Geometrodynamics." *Annals of Physics*, Vol. II (1957).
- Möbius: *Der Barycentrische Calcul*. Leipzig: Barth; 1827.
- Møller, C.: *The Theory of Relativity*. Oxford: Oxford University Press; 1952.
- Nagel, E.: "Einstein's Philosophy of Science." *The Kenyon Review*, Vol. XII (1950).
- : "The Formation of Modern Conceptions of Formal Logic in the Development of Geometry." *Osiris*, Vol. VII (1939).
- : *The Structure of Science*. New York: Harcourt, Brace and World; 1961.
- Newton, I.: *Principia*. Edited by F. Cajori. Berkeley: University of California Press; 1947.
- Northrop, F. S. C.: *The Logic of the Sciences and the Humanities*. New York: The Macmillan Company; 1947.
- : *The Meeting of East and West*. New York: The Macmillan Company; 1946.
- : "Whitehead's Philosophy of Science." *The Philosophy of Alfred North Whitehead*. Edited by P. A. Schilpp. New York: Tudor Publishing Company; 1941.
- Page, L.: *Introduction to Theoretical Physics*. New York: D. Van Nostrand Company, Inc.; 1935.
- Page, L. and Adams, N. I.: *Electrodynamics*. New York: D. Van Nostrand Company; 1940.
- Palter, R. N.: *Whitehead's Philosophy of Science*. Chicago: University of Chicago Press; 1960.
- Panofsky, W. and Phillips, M.: *Classical Electricity and Magnetism*. Cambridge, Mass.: Addison-Wesley Publishing Company; 1955.
- Penrose, O. and Percival, I. C.: "The Direction of Time." *Proceedings of the Physical Society*, Vol. LXXIX (1962).
- Pérard, A.: *Les Mesures Physiques*. Paris: Presses Universitaires de France; 1955.
- Pfeffer, W.: "Untersuchungen über die Entstehung der Schlafbewegungen der Blattoorgane." *Abhandlungen der sächsischen Akademie der Wissenschaften*. Leipzig, Mathematisch-Physicalische Klasse, Vol. XXX (1907).
- Pittendrigh, C. S. and Bruce, V. G.: "Daily Rhythms as Coupled Oscillator Systems and their Relation to Thermoperiodism and Photoperiodism." *Photoperiodism and Related Phenomena in Plants and Animals*. Washington, D. C.: The American Association for the Advancement of Science; 1959.
- : "An Oscillator Model for Biological Clocks." *Rhythmic and Synthetic Processes in Growth*. Edited by D. Rudnick. Princeton: Princeton University Press; 1957.
- Poincaré, H.: *Dernières Pensées*. Paris: Flammarion; 1913.
- : "Des Fondements de la Géométrie, à propos d'un Livre de M. Russell." *Revue de Métaphysique et de Morale*, Vol. VII (1899).
- : *The Foundations of Science*. Lancaster: The Science Press; 1946.

- : *La Mécanique Nouvelle*, cited in R. Dugas: "Henri Poincaré devant les Principes de la Mécanique." *Revue Scientifique*, Vol. LXXXIX (1951).
- : "La Mesure du Temps." *Revue de Métaphysique et de Morale*, Vol. VI (1898).
- : "L'Espace et la Géométrie." *Revue de Métaphysique et de Morale*, Vol. III (1895).
- : *Letzte Gedanken*. Translated by Lichtenecker. Leipzig: Akademische Verlagsgesellschaft; 1913.
- : "The Principles of Mathematical Physics." *The Monist*, Vol. XV (1905). Reprinted in the *Scientific Monthly*, Vol. LXXXII (1956).
- : "Sur les Principes de la Géométrie, Réponse a M. Russell." *Revue de Métaphysique et de Morale*, Vol. VIII (1900).
- : "Sur le problème des trois corps et les equations de la dynamique." *Acta Mathematica*, Vol. XIII (1890).
- Polanyi, M.: "Notes on Professor Grünbaum's Observations." *Current Issues in the Philosophy of Science*. Edited by H. Feigl and G. Maxwell. New York: Holt, Rinehart and Winston; 1961.
- : *Personal Knowledge*. Chicago: Chicago University Press; 1958.
- Popper, K. R.: *The Logic of Scientific Discovery*. London: Hutchinson and Company, Ltd.; 1959.
- : *Nature*, Vol. CLXXVII (1956), Vol. CLXXVIII (1956), Vol. CLXXIX (1957), Vol. CLXXXI (1958).
- Price, D. J.: "The Prehistory of the Clock." *Discovery*, Vol. XVII (1956).
- Putnam, H.: "The Analytic and the Synthetic." *Minnesota Studies in the Philosophy of Science*. Edited by H. Feigl and G. Maxwell. Minneapolis: University of Minnesota Press; 1962. Vol. III.
- : "Three-Valued Logic." *Philosophical Studies*, Vol. VIII (1957).
- Quine, W. V. O.: *From a Logical Point of View*. Second edition. Cambridge: Harvard University Press; 1961.
- Read, J.: *A Direct Entry to Organic Chemistry*. London: Methuen & Company, Ltd.; 1948.
- Reichenbach, H.: *Axiomatik der relativistischen Raum-Zeit-Lehre*. Braunschweig: F. Vieweg & Sons; 1924.
- : "Das Kausalproblem in der Physik." *Naturwissenschaften*, Vol. XIX (1931).
- : "Die Bewegungslehre bei Newton, Leibniz und Huyghens." *Kantstudien*, Vol. XXIX (1924).
- : "Die Kausalstruktur der Welt und der Unterschied von Vergangenheit und Zukunft." *Berichte der Bayerischen Akademie München, Mathematisch-Naturwissenschaftliche Abteilung* (November 1925).
- : *The Direction of Time*. Berkeley: University of California Press; 1956.
- : "Discussion of Dingler's paper." *Physikalische Zeitschrift*, Vol. XXIII (1922).
- : "Kant und die Naturwissenschaft." *Die Naturwissenschaften*, 1933, Vol. XXI.

- : "La signification philosophique du dualisme ondes-corpules." *Louis de Broglie, Physicien et Penseur*. Edited by A. George. Paris: Editions Albin Michel; 1953.
- : "Les fondements logiques de la mécanique des quanta." *Annales de l'Institut Henri Poincaré*, Vol. XIII (1953).
- : *Modern Philosophy of Science*. London: Routledge and Kegan Paul; 1959.
- : *Philosophic Foundations of Quantum Mechanics*. Berkeley: University of California Press; 1948.
- : "The Philosophical Significance of the Theory of Relativity." *Albert Einstein: Philosopher-Scientist*. Edited by P. A. Schilpp. Evanston: Library of Living Philosophers; 1949.
- : *Philosophie der Raum-Zeit-Lehre*. Berlin: Walter de Gruyter & Company; 1928.
- : *The Philosophy of Space and Time*. Dover Publications, Inc.; 1958.
- : "Planetenuhr und Einsteinsche Gleichzeitigkeit." *Zeitschrift für Physik*, Vol. XXXIII (1925).
- : *The Rise of Scientific Philosophy*. Berkeley: University of California; 1951.
- : "Über die physikalischen Konsequenzen der relativistischen Axiomatik." *Zeitschrift für Physik*, Vol. XXXIV (1925).
- : "Ziele und Wege der physikalischen Erkenntnis." *Handbuch der Physik*, Vol. IV (1929).
- Rescher, N.: "On Prediction and Explanation." *British Journal for the Philosophy of Science*, Vol. VIII (1958).
- : "The Stochastic Revolution and The Nature of Scientific Explanation." *Synthese*, Vol. XIV (1962).
- Richter, C. P.: "Biological Clocks in Medicine and Psychiatry: Shock-Phase Hypothesis." *Proceedings of the National Academy of Sciences*, Vol. XLVI (1960).
- Riemann, B.: "On the Hypotheses Which Lie at the Foundations of Geometry." *A Source Book in Mathematics*. Edited by David E. Smith. New York: Dover Publications, Inc.; 1959. Vol. II.
- Robertson, H. P.: "The Geometries of the Thermal and Gravitational Fields." *American Mathematical Monthly*, Vol. LVII (1950).
- : "Geometry as a Branch of Physics." *Albert Einstein: Philosopher-Scientist*. Edited by P. A. Schilpp. Evanston: Library of Living Philosophers; 1949.
- : *Mathematical Reviews*, Vol. XVI (1955).
- : "Postulate versus Observation in the Special Theory of Relativity." *Reviews of Modern Physics*, Vol. XXI (1949).
- Rosenfeld, L.: "On the Foundations of Statistical Thermodynamics." *Acta Physica Polonica*, Vol. XIV (1955).
- Rossi, B. and Hall, D. B.: *Physical Review*, Vol. LIX (1941).
- Rothstein, J.: "Information, Measurement and Quantum Mechanics." *Science*, Vol. CXIV (1951).

- Rougier, L.: *La Philosophie Géométrique de Henri Poincaré*. Paris: F. Alcan; 1920.
- Rush, J. H.: "The Speed of Light." *Scientific American*, Vol. CXCI (1955).
- Russell, B.: *The Foundations of Geometry*. New York: Dover Publications, Inc.; 1956.
- : *Our Knowledge of the External World*. London: George Allen and Unwin, Ltd.; 1926.
- : *The Philosophy of Leibniz*. London: George Allen and Unwin, Ltd.; 1937.
- : *The Principles of Mathematics*. Cambridge: Cambridge University Press; 1903.
- : "Sur les Axiomes de la Géométrie." *Revue de Métaphysique et de Morale*, Vol. VII (1899).
- Saccheri, G.: *Euclides ab omni naevo vindicatus*. Milan: P. A. Montani; 1733.
- Sachs, R. G.: "Can the Direction of Flow of Time Be Determined?" *Science*, Vol. CXL (1963).
- Salecker, H. and Wigner, E. P.: "Quantum Limitations of the Measurement of Space-Time Distances." *The Physical Review*, Vol. CIX (1958).
- Sandage, A.: "Travel Time for Light from Distant Galaxies Related to the Riemannian Curvature of the Universe." *Science*, Vol. CXXXIV (1961).
- Scheffler, I.: "Explanation, Prediction and Abstraction." *British Journal for the Philosophy of Science*, Vol. VII (1957).
- Schlick, M.: "Are Natural Laws Conventions?" *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts, Inc.; 1953.
- : *Grundzüge der Naturphilosophie*. Vienna: Gerold & Company; 1948.
- Scholz, H.: "Eine Topologie der Zeit im Kantischen Sinne." *Dialectica*, Vol. IX (1955).
- Schrödinger, E.: "Irreversibility." *Proceedings of the Royal Irish Academy*, Vol. LIII (1950).
- : "The Spirit of Science." *Spirit and Nature*. Edited by J. Campbell. New York: Pantheon Books; 1954.
- : *What is Life?* New York: The Macmillan Company; 1945.
- Sciama, D. W.: "Inertia." *Scientific American*, Vol. CXCVI (February 1957).
- : "On the Origin of Inertia." *Monthly Notices of the Royal Astronomical Society*, Vol. CXIII (1953).
- Scribner, C.: "Mistranslation of a Passage in Einstein's Original Paper on Relativity." *American Journal of Physics*, Vol. XXXI (1963).
- Scriven, M.: "Comments on Professor Grünbaum's Remarks at the Wesleyan Meeting." *Philosophy of Science*, Vol. XXIX (1962).
- : "Explanation and Prediction in Evolutionary Theory." *Science*, Vol. CXXX (1959).
- : "The Temporal Asymmetry of Explanations and Predictions." *Philosophy of Science*. Edited by Bernard Baumrin. New York: John Wiley and Sons; 1963. Vol. I.

- Sellars, W.: "Time and the World Order." *Minnesota Studies in the Philosophy of Science*, Vol. III. Minneapolis: University of Minnesota Press; 1962.
- Shankland, R. S., McCuskey, S. W., Leone, S. C., and Kuerti, G.: "New Analysis of the Interferometer Observations of Dayton C. Miller." *Reviews of Modern Physics*, Vol. XXVII (1955).
- Shepperd, J. A. H.: "Transitivities of Betweenness and Separation and the Definitions of Betweenness and Separation Groups." *Journal of the London Mathematical Society*, Vol. XXXI (1956).
- Silberstein, L.: *The Theory of Relativity*. New York: The Macmillan Company; 1914.
- Smart, J. J. C.: "Mr. Mayo on Temporal Asymmetry." *Australasian Journal of Philosophy*, Vol. XXXIII (1955).
- : "Critical Study of H. Reichenbach's *The Direction of Time*." *Philosophical Quarterly*, Vol. VIII (1958).
- : "Spatializing Time." *Mind*, Vol. CXIV (1955).
- : "The Temporal Asymmetry of the World." *Analysis*, Vol. XIV (1954).
- Sokolnikoff, I. S.: *Mathematical Theory of Elasticity*. New York: McGraw-Hill Book Company; 1946.
- Sommerfeld, A.: Notes in *The Principle of Relativity*, A Collection of Original Memoirs. New York: Dover Publications, Inc.; 1952. A reprint.
- : *Partial Differential Equations in Physics*. Translated by E. G. Straus. New York: Academic Press, Inc.; 1949.
- Sommerville, D. M. Y.: *The Elements of Non-Euclidean Geometry*. New York: Dover Publications, Inc.; 1958.
- Stanyukovic, K. P.: "On the Increase of Entropy in an Infinite Universe." *Doklady, Akademiia Nauk SSSR, N.S.*, Vol. LXIX (1949).
- Stebbing, L. Susan: *Philosophy and the Physicists*. London: Methuen & Company, Ltd.; 1937.
- Stille, U.: *Messen und Rechnen in der Physik*. Braunschweig: Vieweg; 1955.
- Struik, D. J.: *Classical Differential Geometry*. Cambridge: Addison-Wesley Publishing Company, Inc.; 1950.
- Tamm, I. E.: "General Characteristics of Vavilov-Cherenkov Radiation." *Science*, Vol. CXXXI (1960).
- : "Radiation of Particles With Speeds Greater Than that of Light." *American Scientist*, Vol. XLVII (1959).
- Tarski, A.: "What is Elementary Geometry?" *The Axiomatic Method*. Edited by L. Henkin, P. Suppes, and A. Tarski. Amsterdam: North Holland Publishing Company; 1959.
- Taub, A. H.: "Empty Space-Times Admitting a Three Parameter Group of Motions." *Annals of Mathematics*, Vol. LIII (1951).
- Taylor, G. I.: "Tidal Friction in the Irish Sea." *Philosophical Transactions of the Royal Society Academy*, Vol. CCXX (1920).
- Taylor, W. B.: *The Meaning of Time in Science and Daily Life*. Doctoral Dissertation. Los Angeles: University of California at Los Angeles; 1953.

- Ter Haar, D.: "Foundations of Statistical Mechanics." *Reviews of Modern Physics*, Vol. XXVII (1955).
- Thomas, T. Y.: *The Differential Invariants of Generalized Spaces*. Cambridge: Cambridge University Press; 1934.
- Timoshenko, S. and Goodier, J. N.: *Theory of Elasticity*. New York: McGraw-Hill Book Company; 1951.
- Tolman, R. C.: *Relativity, Thermodynamics and Cosmology*. Oxford: Oxford University Press; 1934.
- : *The Principles of Statistical Mechanics*. Oxford: Oxford University Press; 1938.
- Trocheris, M. G.: "Electrodynamics in a Rotating Frame of Reference." *Philosophical Magazine*, Vol. XL (1949).
- Truesdell, C.: "Ergodic Theory in Classical Statistical Mechanics." Edited by P. Caldivole. *Ergodic Theories*. New York: Academic Press; 1961.
- Tannery, P.: "Le Concept Scientifique du Continu: Zenon d'Élée et Georg Cantor." *Revue Philosophique*, Vol. XX, No. 2 (1885).
- Veblen, O. and Whitehead, J. H. C.: *The Foundations of Differential Geometry*. (Cambridge Tracts in Mathematics and Mathematical Physics, No. 29.) Cambridge: Cambridge University Press; 1932.
- von Fritz, K.: "The Discovery of Incommensurability by Hippasus of Metapontum." *Annals of Mathematics*, Vol. XLVI (1945).
- von Helmholtz, H.: *Schriften zur Erkenntnistheorie*. Edited by P. Hertz and M. Schlick. Berlin: J. Springer; 1921.
- von Laue, Max: *Die Relativitätstheorie*. Braunschweig: F. Vieweg; 1952. Vol. I.
- : "A. Einstein und die Relativitätstheorie." *Naturwissenschaften*, Vol. XLIII (1956).
- von Neumann, J.: *Mathematische Grundlagen der Quantenmechanik*. Berlin: J. Springer; 1932. Translated by R. T. Beyer. Princeton: Princeton University Press; 1955.
- von Weizsäcker, C. F.: "Der zweite Hauptsatz und der Unterschied von Vergangenheit und Zukunft." *Annalen der Physik*, Vol. XXXVI (1939).
- Waismann, F.: *Introduction to Mathematical Thinking*. New York: F. Ungar Publishing Company; 1951.
- Walker, A. G.: "Axioms for Cosmology." *The Axiomatic Method*. Edited by L. Henkin, P. Suppes, and A. Tarski. Amsterdam: North Holland Publishing Company; 1959.
- Ward, F. A. B.: *Time Measurement*. Fourth edition. London: Royal Stationery Office; 1958. Part I.
- Watanabe, M. S.: "Le Concept de Temps en Physique Moderne et la Durée Pure de Bergson." *Revue de Métaphysique et de Morale*, Vol. LVI (1951).
- : *Le Deuxième Théorème de la Thermodynamique et la Mécanique Ondulatoire*. Paris: Hermann & Cie.; 1935.
- : "Réversibilité contre Irréversibilité en Physique Quantique." *Louis de Broglie, Physicien et Penseur*. Paris: Albin Michel; 1953.
- : "Symmetry of Physical Laws. Part III, Prediction and Retrodiction." *Reviews of Modern Physics*, Vol. XXVII (1955).

- : "Über die Anwendung Thermodynamischer Begriffe auf den Normalzustand des Atomkerns." *Zeitschrift für Physik*, Vol. CXIII (1939).
- Weber, J.: "Phase, Group, and Signal Velocity." *American Journal of Physics*, Vol. XXII (1954).
- Weinstein, B.: *Handbuch der Physikalischen Massbestimmungen*. Berlin: J. Springer; 1886, Vol. I, and 1888, Vol. II.
- Weiss, P.: *Reality*. New York: Peter Smith; 1949.
- Weyl, H.: "50 Jahre Relativitätstheorie." *Naturwissenschaften*, Vol. XXXVIII (1951).
- : *Philosophy of Mathematics and Natural Science*. Princeton: Princeton University Press; 1949.
- : *Space-Time-Matter*. New York: Dover Publications, Inc.; 1950.
- Wheeler, J. A.: "Curved Empty Space-Time as the Building Material of the Physical World: An Assessment." Edited by E. Nagel, P. Suppes, and A. Tarski. *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*. Stanford: Stanford University Press; 1962.
- : "Mach's Principle as a Boundary Condition for Einstein's Field Equations and as a Central Part of the 'Plan' of General Relativity." A Report given at the Conference on Relativistic Theories of Gravitation, Warsaw, July 1962.
- : "The Universe in the Light of General Relativity." *The Monist*, Vol. XLVII, No. 1 (1962).
- Whitehead, A. N.: *The Concept of Nature*. Cambridge: Cambridge University Press; 1926.
- : *Essays in Science and Philosophy*. New York: The Philosophical Library; 1947.
- : *The Principles of Natural Knowledge*. Cambridge: Cambridge University Press; 1955.
- : *The Principle of Relativity*. Cambridge: Cambridge University Press; 1922.
- : *Process and Reality*. New York: The Macmillan Company; 1929.
- Whitrow, G. J.: *The Natural Philosophy of Time*. London: Thomas Nelson & Sons, Ltd.; 1961.
- Whittaker, E. T.: *From Euclid to Eddington*. London: Cambridge University Press; 1949.
- : "G. F. Fitzgerald." *Scientific American*, Vol. CLXXXIX (November 1953).
- : *A History of the Theories of Aether and Electricity*. London: Thomas Nelson & Sons, Ltd.; 1953.
- Whyte, L. L.: "Light Signal Kinematics." *British Journal for the Philosophy of Science*, Vol. IV (1953).
- : "One-Way Processes in Physics and Biophysics." *British Journal for the Philosophy of Science*, Vol. VI (1955).
- Wiener, N.: *Cybernetics*. New York: John Wiley & Sons, Inc.; 1948.
- Wigner, E. P.: "Relativistic Invariance and Quantum Phenomena." *Reviews of Modern Physics*, Vol. XXIX (1957).

- Yanase, M. M.: "Reversibilität und Irreversibilität in der Physik." *Annals of the Japan Association for Philosophy of Science*, Vol. I (1957).
- Zawirski, Z.: *L'Évolution de la Notion du Temps*. Cracow: Librairie Gebethner & Wolff; 1936.
- Zermelo, E.: "Über einen Satz der Dynamik und der mechanischen Wärmetheorie." *Wiedemannsche Annalen (Annalen der Physik und Chemie)*, Vol. LVII (1896).
- Zilsel, E.: "Über die Asymmetrie der Kausalität und die Einsinnigkeit der Zeit." *Naturwissenschaften*, Vol. XV (1927).

PART IV

Supplementary Studies

1964-1973

SPACE, TIME AND FALSIFIABILITY CRITICAL EXPOSITION AND REPLY TO "A PANEL DISCUSSION OF GRÜNBAUM'S PHILOSOPHY OF SCIENCE"

PART I*

Prompted by the "Panel Discussion of Grünbaum's Philosophy of Science" (*Philosophy of Science* 36, December, 1969) and other recent literature, this essay ranges over major issues in the philosophy of space, time and space-time as well as over problems in the logic of ascertaining the falsity of a scientific hypothesis.

The author's philosophy of geometry has recently been challenged along three main distinct lines as follows: (i) The Panel article by G. J. Massey calls for a more precise and more coherent account of the Riemannian conception of an intrinsic as opposed to an extrinsic metric, which the author has invoked as his basis for the distinction between non-conventional and convention-laden ascriptions of metrical equality and inequality; (ii) the latter distinction has been claimed to suffer from the liabilities of the so-called "standard conception" of scientific theories [36]; and (iii) pursuant to H. Putnam's "An Examination of Grünbaum's Philosophy of Geometry" [56], J. Earman [16, 17] and R. Swinburne [65] have contended that the difference between intrinsic and extrinsic metrics is scientifically unilluminating, and that the associated distinction between non-conventional and convention-laden metrical comparisons does not have the kind of relevance to extant scientific theories that the author has claimed for it.

The essay consists of two installments. The present installment, comprising the **Introduction** and **Part A**, is devoted to the clarification, correction and further development of the author's prior writings on the philosophy of geometry. Its main objective is constructive elaboration rather than offering polemics. But rebuttals to Earman, [16, 17], Swinburne [65] and Demopoulos [13] are included, because their inclusion conduces to clarity in giving the new exposition. **Part B** is to appear in a subsequent issue and will be devoted to *replies* to critiques contained in the Panel Discussion and in other recent literature. It will range over issues in the philosophy of geometry and in the logic of ascertaining the falsity of a scientific hypothesis.

By way of merely elliptical anticipation of much more precise statements given in **Part A**, section 3(ii), the Introduction *dissociates* the notion of *convention-ladenness* developed in **Part A** from the *quite different* notion integral to the so-called "standard conception" of scientific theories. Thereby, the Introduction prepares the ground for seeing, as a corollary to **Part A**, section 3(ii), that the notion advocated in the present essay has nothing to fear from the following fact, noted by C. G. Hempel ([36]; cf. also his 1970 Carus Lectures): "even though a sentence may originally be introduced as true by stipulation, it soon joins the club of all other member-statements of the theory and becomes subject to revision in response to further empirical findings and theoretical developments."

Part A, which begins with a fairly detailed table of contents, endeavors to meet the aforementioned three-fold challenge to the author's philosophy of geometry.

* **PART II** will be published in *Philosophy of Science*.

Massey's call for the provision of clear and detailed characterizations of intrinsic and extrinsic metrics is answered with the invaluable aid rendered personally by Massey himself. These characterizations are shown to permit a precise explication of the portions of Riemann's Inaugural Dissertation relevant to (1) Riemann's brilliant anticipation of Einstein's *dynamical* conception of physical geometry, and (2) the author's philosophical characterization of the metrics and geometries of space, time, and space-time (section 2(c)).

A byproduct of the analysis is to raise two major philosophical doubts concerning Clifford's sketch of a so-called "space-theory of matter" as elaborated in J. A. Wheeler's relativistic geometrodynamics. That theory's vision of understanding matter as a manifestation of empty curved space is questioned in regard to (1) the existence of an *intra*-geometrodynamical basis for individuating the metrically homogeneous world points of its space-time manifold (section 1(a)), and (2) the compatibility of the theory with the Riemannian metrical philosophy apparently espoused by its proponents (section 2(c) (i)).

See
Append.
§42

Introduction. During the past year or two, several friends and colleagues, including two former students, Bas van Fraassen and Philip Quinn, have done me the honor of devoting a number of articles to some of my earlier work. Ranging over the falsifiability of scientific hypotheses and certain topics in the philosophy of space and time, a number of these articles appeared as a Panel in the December 1969 issue of this Journal. The Editor-in-Chief kindly invited me to write an omnibus Reply. And he encouraged me to comment also on any other published criticisms of my views with which I wish to deal and have not dealt before.¹ I am gratefully availing myself of this welcome opportunity in the pages that follow. But my main objective is *not* polemical. Instead, I am principally concerned to provide a further development and clarification of ideas, as rightly urged by Massey. And in the course of doing so, I hope to correct some errors and mis-statements of mine.

I shall present the discussion in two main parts as follows: **A.** Criteria for Intrinsicness vs. Extrinsicness of Metrics and of Relations on Manifolds, and **B.** Reply to Critiques: I. Issues in the Philosophy of Space and Time, and II. Problems in the Logic of Ascertaining the Falsity of a Scientific Hypothesis. **Part A** is appearing as the first installment.² A detailed table of contents of it is given at the end of this Introduction.

Although replies to critiques are generally deferred to **Part B**, exceptions were made when the inclusion of a reply was directly helpful in giving the new exposition to which **Part A** is devoted.

A portion of **Part A** below will be devoted to a clarification of the concept of an

¹ The most recent earlier replies which I have published are contained in [28], Ch. III and [31], [29], [28], Ch. II, Section 1, [30], [32] and [25].

² *Acknowledgments:* Apart from the specific acknowledgments to Gerald Massey within the text, I wish to emphasize my *immense* debt to him, since this installment could not have materialized in anything like its present form without his constant help and advice. He rendered invaluable assistance to me on matters of substance and style, saved me from various blunders and offered line by line comments on the entire manuscript. I am likewise most grateful to Allen Janis who is perennially generous with his time in giving me scientific information on matters of physics. Furthermore, I am indebted to Nuel Belnap, Alberto Coffa, Philip Quinn, Bas van Fraassen and Wilfrid Sellars for helpful discussions of some of the pertinent materials.

Finally, I wish to thank the National Science Foundation for the support of research on which this paper is based.

extrinsic vs. an intrinsic metric. And one of the philosophical results will be to articulate a good deal more clearly than heretofore (with respect to nontrivial metrics) the sense in which *the assumptions made about the world* by specific theories of space and time imply the presence or absence of importantly conventional ingredients in true descriptive uses of the intensionally preempted terms “congruent” and “incongruent.” In this context, ascriptions of congruence or incongruence are to be construed as ascriptions of a relational *property* in the intensional sense and not as mere extensional assertions or denials of membership in a set of ordered pairs or in an equivalence class. Thus, it will hopefully become quite clear in what sense the ascriptions of metrical equality and inequality generated by nontrivial metrics in specific theories of space and time—and hence the corresponding geometries of space, time and space-time—*do* or *do not* involve importantly conventional ingredients. To accord with the language used by such foundational thinkers as Riemann and Russell and also for brevity, I shall often say that metrical or congruence (and incongruence) *relations* have a certain kind of philosophical status, but I shall thereby *mean* that the corresponding metrical *ascriptions* have that status. Thus, I may say that a congruence relation in the class of spatial intervals of a continuous physical 3-space involves an important conventional ingredient. But this statement should *not* be understood literally as pertaining to a two-termed relation-in-extension among spatial intervals. I refer to another, very recent publication [32] for my correspondingly detailed elucidation of the important conventional ingredient in ascriptions of distant simultaneity, made by Einstein’s theory of relativity.

With respect to certain particular physical theories, I asserted in earlier writings the conventional status of congruence in space, time and space-time, and of simultaneity. It has been mistakenly thought that these particular claims of mine can be construed as mere instances—pertaining to the particular theoretical terms “spatially congruent”, “temporally congruent,” “spatio-temporally congruent,” and “distantly simultaneous”—of a quite *general* thesis which *some* philosophers have asserted alike for any and *every* theoretical term and for *every* physical theory. Yet I had explained, and will argue further in the sequel for the case of congruence, why I would *deny* the conventionality of the corresponding ascriptions of congruence and simultaneity, *if* certain *other* physical theories were held to be true. It has been furthermore erroneously supposed that my particular contentions about the ingredience of nontrivial conventions are predicated on a conception that makes insufficient allowance for the *multiple criteria* character of the theoretical terms employed by scientific theories. In the main body of this essay, I hope to strengthen my previously published rebuttals to both of these mistaken suppositions.³ But it will be useful at this point to call attention to some key misunderstandings with which I have not dealt sufficiently before and which underly the second of these mistaken suppositions.

³ I refer to my [28], Ch. III, Sections 2.10–2.11, 3 and 4 for a critique of the first, and to my [24], pp. 48–56 and [28], Ch. II, Section I, as well as pp. 217–218 and pp. 270–281 for my objections to the second.

We shall see in section 2(a), subsection 2, that in the technical sense of "congruent," congruence is merely one particular *species* of metrical equality. But pending our statement of the differences among some of the major species of metrical equality in section 2(a), I shall use the term "congruence" in this Introduction in a nontechnical sense to refer indifferently to any species of metrical equality.

Assume that in the context of a certain physical theory, the congruence class or partition in a stated class of spatial intervals is specified alike by each member of a cluster of physical criteria of spatial congruence. In that event, these several criteria are concordant or interchangeable as a matter of empirical *fact*. Thus, according to the theory of relativity, in *any* inertial system, spatial intervals which are equal in the metric based on rigid rods are also equal in the metric furnished by the round trip times of light. Now suppose further, merely for simplicity, that there *were* only these two criteria of spatial congruence in that theory and that I then reiterate two claims from a 1968 book ([28], pp. 156–157, and 217–218), which I shall state after first entering the following *caveat*: the language in which they are stated is elliptical and can be misleading. Its proper construal is set forth in section 3, and I use it *preliminarily* in this Introduction only in the interest of continuity with locutions in the long-established prior literature familiar to many readers. The two claims are: (1) Given the continuity postulated by the theory of relativity for the physical 3-space of each inertial frame, the self-congruence under transport of some *one* of these two different kinds of standards of spatial congruence is conventional, i.e. its qualifying as a standard to begin with is conventional, and either kind of standard may be conventionally established as self-congruent under transport to begin with. (2) Any appeal to the *concordance* of the two standards to certify the self-congruence under transport of one of them *as a fact* requires the *conventional* assertion of the spatial self-congruence of the other. For their mere concordance does *not* suffice to establish that either or both of them is spatially self-congruent under transport. And precisely the same is true of such concordance as obtains among a whole plethora of kinds of standards! As Clifford put it so lucidly:

... The reader will probably have observed that we have defined length or distance by means of a measure which can be carried about *without changing its length*. But how then is this property of the measure to be tested? We may carry about a yard measure in the form of a stick, to test our tape with; but all we can prove in that way is that the two things are always of the same length when they are in the same place; not that this length is unaltered. . . .

Is it possible, however, that lengths do really change by mere moving about, without our knowing it?

Whoever likes to meditate seriously upon this question will find that it is wholly devoid of meaning ([9], pp. 49–50).

My claims here make it explicit that the concordance of two or more kinds of standards of congruence, no less than that of the members of one kind, is a matter of fact (empirical law), as opposed to being stipulational. By the same token, I affirm, of course, that the discordance or non-interchangeability of two or more standards of congruence is a matter of fact. Thus, I have emphasized ([24], p. 67) the factual character of the discordance between the time congruence based on the earth's rotational motion, (the "diurnal" or "sidereal" time congruence), and the

refined “Newtonian” time congruence based on the requirement of the validity of Newton’s laws of motion, as modified by the very small corrective terms expressing the relativistic motion of the perihelia. In that connection, I likewise noted the discordance between the diurnal (sidereal) time congruence and the *unrefined* Newtonian one implicit in the unmodified statement of Newton’s laws. And my view clearly holds it to be factual and *nonstipulational* that in the context of the original nineteenth-century aether theory of light, spatial intervals in *moving* inertial systems which are equal in the rigid rod metric will generally be *unequal* in the metric based on the travel times of light. For philosophical reasons that are set forth in Section 3(ii) below, I do NOT use the terms “stipulational” and “conventional” interchangeably; instead, my *special* usage of the term “conventional” is spelled out and justified there.

Hence my assertion of the ingreience of a convention is coupled with a claim of factuality that prompts me to oppose the following contention as *incorrect*: When a physical theory provides two or more concordant test conditions for the applicability of a particular term like “spatially congruent” by means of a conjunction of sentences, then each and *every* one of these sentences is true by stipulation such that no new facts can falsify their conjunction.⁴ Plainly, I reject this contention. For example, I certainly allow that the special theory of relativity might have to be amended to state, as has recently been suggested, that in inertial frames the travel times of light are generally *not* interchangeable with rigid rods as a spatial congruence standard after all. According to this suggested emendation, light does *not* have the constant speed c independently of its frequency and has positive rather than zero rest-mass. And on that assumption, light does *not* generally take equal times to traverse distances which are equal in the metric based on rigid rods.⁵ If tests on light from pulsars should bear out a frequency-dependence of its speed, then the travel times of light (“radar ranging”) and rigid rods would not furnish concordant criteria of spatial congruence in inertial frames, although the travel times of any “particles” of *zero* rest mass and rigid rods could still be held to do so.

Analogously, my claims obviously make logical provision for the pragmatically-motivated abandonment of the initial diurnal criterion of time congruence by the conventional adoption of the different Newtonian one, a switch which then issues in the conclusion that sentences that were true assertions of temporal equality with respect to the diurnal metric become *false* with respect to the Newtonian one.⁶ Moreover, I have previously noted the following à propos of temporal congruence ([24], pp. 55–56 and 67): After the earth was qualified as a standard clock by conventionally declaring its axial rotation to be uniform, observations of the moon’s

⁴ This kind of stipulationalist assertion, which is implicit in one of the prevalent conceptions of scientific theories, has also been rightly criticized by Hempel ([36], Section 6.)

⁵ For a statement of precisely this kind of proposed modification of the special theory of relativity, see [22].

⁶ But I argued in Ch. 2 of [24] that though these two criteria of time congruence are indeed incompatible, they can be used to furnish equivalent descriptions of the factual content of Newtonian mechanics: A precise statement of the sense in which this claim of equivalence of descriptions is to be understood is given in Part A, section 3(ii).

motion occasioned developments illustrating Hempel's remark "that certain laws or theoretical principles originally based on evidence that includes the readings of standard clocks may call for the verdict that those clocks do not mark off strictly equal time intervals" [36]. Hence I agree with Hempel that numerous and important episodes in the history of science do furnish illustrations of his claim that "even though a sentence may originally be introduced as true by stipulation, it soon joins the club of all other member-statements of the theory and becomes subject to revision in response to further empirical findings and theoretical developments" [36]. Thus construed, this statement of Hempel's is compatible with my answer to the following question: If a given kind of clock *C* is to be disavowed as nonuniform (nonisochronous) after an initial convention that it is a self-congruent time standard, then does the fact of *C*'s discordance with another kind of time standard *U* confer *factual* temporal truth on the assertion of *C*'s nonuniformity? I reply that, if the instants of time are a mathematical continuum, then *C* can be indicted as non-uniform in virtue of its discordance with *U* only relatively to the conventional affirmation of the isochronism (self-congruence) of the *new*, alternative standard *U*.⁷ For certain facts may render it highly inconvenient to employ *C* as a conventional standard of isochronism in the description of nature without thereby detracting from the conventionality of the standard *U* that is used in its stead. Nor is the conventionality of *U* as a standard impugned by *U*'s subsequent survival as a standard for an extended period amid much other scientific change. *Mutatis mutandis*, the same remarks apply to any two possible standards of spatial congruence and to the longevity of one of them.

From 1889 to 1960, the platinum-iridium rod served as the standard meter amid much scientific change. And in 1960, there was a switch to the as of then concordant (equivalent) standard furnished by 1,650,763.73 wave lengths of orange-red light of the gas krypton 86 ([48], pp. 46–47). But according to speculations by Dirac and Dicke, it may turn out that these initially concordant length standards of atomic constitution may become discordant with time ([48], pp. 47–48). Indeed, R. F. Marzke and J. A. Wheeler have suggested ([48], p. 48) that neither of these two standards *may* turn out to yield values of proper distances compatible with classical general relativity theory. Such a development would be analogous to the aforementioned incompatibility of the diurnal time congruence with the durational equalities required for the validity of Newton's laws of motion. For this reason, these authors have argued that classical general relativity ought to and indeed can *dispense* with both the meter bar and the krypton meter, as well as with any other length or duration standards of atomic constitution ([48], pp. 40–64).

Returning to the time congruence implicit in the statement of Newton's (unmodified) laws, it is worth noting that quartz clocks, and apparently also molecular

⁷ Thus, suppose that in E. A. Milne's theory, which postulates it to be a fact that astronomical clocks and atomic ones furnish discordant time congruences (cf. [24], pp. 22–23), the former clocks were rejected as nonisochronous in favor of the latter. Then I maintain that the factual noninterchangeability of the two kinds of clocks cannot confer factual temporal truth on the claim that astronomical clocks are nonisochronous, unless the isochronism of atomic clocks is first asserted conventionally.

and atomic clocks, advance *nonuniformly* with respect to that Newtonian time congruence no less than sidereal (diurnal) and pendulum clocks do ([3], p. 16). And this poses the still unresolved question of how the rates of these various kinds of clocks are related to the proper time congruences furnished by theoretical standard clocks that satisfy the demands of the theory of relativity (cf. [3], p. 22).

Though the evidence from the history of science thus does exemplify Hempel's statement, that evidence seems to me not to bear out as well an apparently stronger statement by Quine which Hempel goes on to quote in his very next sentence, writing: "As Quine has said, 'conventionality is a passing trait significant at the moving front of science, but useless in classifying the sentences behind the lines.' " To my mind, this statement is misleading by omitting important considerations. For, in the first place, I take it that shortly before 1960 the sentences which stipulated the self-congruence of the differentially unperturbed Paris meter bar under transport and the (relativistically refined) Newtonian time congruence mentioned above belonged to Quine's category "behind the lines." Yet shortly before 1960 these sentences qualified as conventional just as significantly as any sentence "at the moving front of science" that does so qualify initially in Quine's view. And whatever the metrical criteria used at particular times, these are not the less ineluctably conventional in a theory asserting the continuity of space and time, just because empirical findings may have played a role in prompting scientists to give up *particular* criteria at particular times. More fundamentally, I have two reasons for rejecting as misleading the thesis of the *mere* ephemerality of conventional ingredients in the specified ascriptions of metrical equality made by certain kinds of scientific theories: (i) The cited remarkable *longevity* of the conventional self-congruence of a rigid rod, for example, and (ii) As long as continuity is postulated as a feature of the manifolds of space, time, and space-time, there is an ineluctable conventionality of whatever standards of spatial, temporal, or spatio-temporal self-congruence are being used at any given time, which makes for the *long-term* presence of some self-congruence convention or other in physical theory, even though *particular* standards are being supplanted by others. For the metrical equalities and inequalities generated by any nontrivial metric defined on the intervals of the specified kinds of manifolds are presumed to be devoid of an intrinsic basis in the sense of Riemann (see section 2(c)). And given that Riemannian assumption, these metrical ascriptions involve an important conventional ingredient. Moreover, the conventional status of that ingredient is unaffected by the fact that the extrinsic basis of the metric is furnished by a whole plethora of concordant kinds of physical standards rather than by merely one kind of standard (e.g. meter sticks).

Having said this, I hope nonetheless that the sequel will make clearer than I have made heretofore the full extent of my rejection of both the substance and the spirit of *unbridled* conventionalism. Why devote attention to pointing out the non-trivially conventional elements ingredient in stated ascriptions made by designated scientific theories? One of my grounds for doing so is to exhibit anew that in the context of the particular given theories, these special elements have the interesting status of being conventional only for the following reason: The theories in question

presume certain particular states of affairs that are not themselves conventional but have the character of being matters of spatial, temporal or spatio-temporal fact! And given an alternative theory making suitably different factual assumptions as to these states of affairs, the counterparts of ascriptions that were convention-laden in the original theory will no longer be so in the alternative theory.

Pending the further critique of unbridled conventionalism to be given later in this essay, let me just mention here some illustrations of my disagreements with it: (i) the anti-conventionalist account, to be found in **Part A**, sections 2 and 3, of the ontological and epistemic status of certain metrical equalities in discrete (quantized) physical space, (ii) my recent arguments ([32], pp. 19–24) for the completely *non*-conventional character of *Newtonian* simultaneity, and (iii) the critiques which I have offered of the presently influential conventionalist doctrine that any one component hypothesis of a scientific theory may be held to be true come what may ([25] and [30]).

Part A. Criteria for Intrinsicness vs. Extrinsicness of Metrics and of Relations on Manifolds

CONTENTS

1. Singly and Multiply Extended Manifolds	458
(a) "Spaces": Manifolds whose elements are metrically homogeneous with one another	458
(b) Manifolds whose elements are intrinsically inhomogeneous with one another: <i>I</i> -manifolds with respect to a set <i>C</i> of intrinsic properties	467
2. Intrinsicness vs. Extrinsicness of Metrics, Metrical Equalities, and Congruences	468
(a) Distance-metrics, measure-metrics, and metrics of a Riemannian space	468
1. Distance-metrics ("D-metrics")	468
2. Measure-metrics ("M-metrics")	469
3. Riemann-metrics ("R-metrics")	471
(b) Criteria for the nontriviality of several species of metrics	474
1. General remarks	474
2. <i>M</i> -metrics	475
3. Length functions ("L-metrics")	488
4. <i>D</i> -metrics	490
5. Indefinite <i>R</i> -metrics of relativity theory	492
6. Conclusions on nontriviality	492
(c) Intrinsicness of a metric to the intervals of a manifold, to ordered pairs of elements of a manifold, and intrinsicness to the manifold as a whole	495
(i) Preliminary remarks on the underlying conception; its bearing on a space-theory of matter	495
(ii) Statement of criteria of intrinsicness and extrinsicness for metrics defined on intervals (" <i>m</i> -metrics")	505
(iii) Intrinsicness and extrinsicness of <i>D</i> -metrics	510
(iv) Intrinsic and extrinsic metrics in several varieties of manifolds	512
1. The real number continuum	512
2. The absolute rational numbers	517
3. The natural numbers	517
4. Discrete <i>P</i> - (or <i>T</i> -) manifolds	518
(α) Bilaterally boundless discrete denumerable <i>P</i> - (or <i>T</i> -) manifolds	519
(β) Unilaterally bounded and bilaterally bounded discrete <i>P</i> - (or <i>T</i> -) manifolds	523
5. Continuous <i>P</i> - (or <i>T</i> -) spaces	526
Statement of RMH (generalized)	527
6. Dense denumerable <i>P</i> - (or <i>T</i> -) spaces	531
7. Time-dependent spatial metrics and the dynamical conception of the space-time metric	535
3. What are the Logical Connections, if any, between Alternative Metrizability, Intrinsic Metric Amorphousness, and the Convention-ladenness of Metrical Comparisons?	547
(i) Is there any connection between alternative metrizability and intrinsic metric amorphousness?	547
(ii) What is the connection between the convention-ladenness of metrical comparisons and the extrinsicness of their bases of comparison?	556
(iii) Is normative alternative metrizability either sufficient or necessary for the convention-ladenness of metrical comparisons?	561
4. Intrinsicness and Extrinsicness of a Relation on a Manifold	563

1. Singly and Multiply Extended Manifolds

(a) *Spaces: Manifolds whose elements are metrically homogeneous with one another*

In his *Foundations of Geometry*, Bertrand Russell wrote ([63], pp. 66–67): “With space, . . . , if we seek for elements, we can find none except points, and no analysis of a point will find magnitudes inherent in it. . . . The magnitudes which space deals with, . . . , are relations between points. . . . There is no inherent quality in a single point, as there is in a single colour, by which it can be quantitatively distinguished from another.”

Can the points of physical space be *homogeneous* with respect to all qualities relevant to the spatial extension of their associated singletons, and yet be *individuated* by monadic and/or polyadic physical attributes or by other means? The answer is “yes,” only if those attributes that serve to individuate the points are not deemed relevant to the spatial extension of the singletons to which they belong. Thus, Russell seems to consider the points of space as given individuals and claims that they are all alike at least in the respect that the singletons to which they belong have the same spatial extension. Hence, he is asserting the *metrical equality* of the singletons of the points of space. And this claim of *metrical homogeneity* leaves open whether physical space is quantized, continuous or something else. A discussion of the warrant for my countenancing Russell’s ascription of metrical equality to *singletons* of physical space will be deferred to sections 2(b) and 2(c), and 3 below. There we shall also discuss some “pathological” measure functions on *discrete* physical space which do *not* assign equal measures to the singletons of that manifold. But since these measure functions are exceptional, it will simplify our exposition to postpone discussion of them until sections 2(b)–2(c). And hence assertions of metrical homogeneity which I shall make concerning discrete physical space in the ensuing prior sections are to be understood as not begging the question in regard to the exceptional “pathological” measures on that kind of manifold.

Mutatis mutandis, we can make the same statements of metrical homogeneity about physical time and space-time. Hence we can also say in the specified metrical sense that qua elements of physical time, instants are all alike, and qua elements of Einstein’s space-time, punctal events are metrically homogeneous with one another.

While the elements of a manifold may be homogeneous with respect to the metrical extension of their associated singletons, they may not be homogeneous in other respects. Thus there are respects in which the individual natural numbers, construed as the usual arithmetic objects, are *not* homogeneous with one another *qua arithmetic elements*, any more than the colors in the spectrum of physical optics mentioned by Russell. And similarly, the individual rational numbers are not all alike. Nor are the elements of the real number continuum. For in each of these number manifolds, the individual numbers are *non-homogeneous* with respect to the set of familiar arithmetic properties that come within the purview of the familiar arithmetic theory of these manifolds. Indeed arithmetic as a discipline does take cognizance of differences between its elements that can serve to *individuate* each and every one of them. Hence we can say that the elements of our arithmetic

number manifolds are not merely non-homogeneous but inhomogeneous with respect to their familiar arithmetic properties. Yet the singletons of the manifold of real numbers, for example, would be *metrically* homogeneous with respect to a length function L given by $L[a, b] = |b - a|$, where a and b are any real numbers, $[a, b]$ is the closed interval determined by them, and $|b - a|$ is the absolute value of the arithmetic difference between a and b .

I shall use the term "space" to refer to any manifold on which a measure function assigning the same value to each singleton has been defined, i.e. to refer to any manifold whose singletons are *homogeneous with respect to metrical extension*. More precisely, by the term "space," I shall mean an ordered pair consisting of a manifold together with a measure function that assigns the same value to all the singletons of the manifold. (In a wider sense, I shall also occasionally speak of a manifold on which a distance function has been defined as being a "space" with respect to that kind of gauge function. For a distance function always assigns *the same value* zero to the ordered pair (a, a) as the distance of a point a from itself.) This usage allows one and the same manifold to qualify as a "space" with respect to some measure functions while not so qualifying with respect to others. But it likewise allows that the ascription of metrical homogeneity to the singletons of a particular manifold be grounded in the particular homogeneous constitution of its elements, as in the case of the points of physical space. At times, clumsy wording will be avoided by referring to the elements of the manifold rather than to their associated singletons as *metrically* homogeneous, although this is not literally correct, strictly speaking.

I noted by reference to a length function on the manifold of real numbers that when speaking of a manifold as "metrically homogeneous," I mean that its singletons are metrically homogeneous with one another. But I also pointed out that metrical homogeneity does *not* entail that the individual members of the singletons have the same constitution qua elements of the manifold, i.e. metrical homogeneity does *not* require the elements of the manifold to be homogeneous with respect to properties *other than* the measures of their associated singletons. Hence the metrical homogeneity of the singletons of a manifold must *not* be identified with such homogeneity as may obtain in the constitution of the elements of the manifold themselves. Nonetheless, in particular cases, the metrical homogeneity of the singletons may be grounded in the constitutional homogeneity of the elements themselves. I was employing the latter constitutional sense of homogeneity when I spoke of the elements of a manifold as being homogeneous in earlier publications ([23], p. 431 and [24], pp. 16–17). Indeed, the difference between the two senses of homogeneity makes itself felt further in our example of a length function on the real numbers in the following way: the very *differences* in the arithmetic constitution of the various real number elements themselves are used to generate the non-trivial length metric $|b - a|$ on that number manifold, although this same metric renders the singletons of that manifold metrically homogeneous with one another.

I shall speak of a manifold as *inhomogeneous*, as distinct from merely *nonhomogeneous*, with respect to a set C of properties (both monadic and polyadic), iff *all* the elements of the manifold are individuated by C . For brevity, I shall refer to a

manifold that is inhomogeneous with respect to C as an " I -manifold" with respect to C . It is clear from the above in what sense our three arithmetic number manifolds are each inhomogeneous or I -manifolds with respect to a set of familiar arithmetic properties.

In the present Section (a), my concern is with *spaces*. But I shall distinguish physical space, physical time, and physical space-time from one another by prefixes as follows: P -space (physical space), T -space (time), and PT -space (space-time). And each of these can, of course, be subdivided into structurally different kinds.

Consider the set S of intuitively singly extended P -spaces containing discrete P -space, dense denumerable P -space, the linear continuous P -space whose points form a topological Euclidean open line, and a *closed line* in a P -manifold that forms a Euclidean topological plane. Our set S of P -spaces contains both zero-dimensional and one-dimensional members, as dimensionality is defined in the present-day topological theory of dimension [41]. On the other hand, the "quantized" P -space whose elements are the *tiles* that might cover the surface of a 4-cornered table is intuitively *not* singly-extended, although it is of finite cardinality and is no less zero-dimensional than singly-extended discrete P -space. Hence the attempt to explicate an intuitively singly extended space as one which is either zero-dimensional or one-dimensional would fail by being too inclusive.

In his contribution to the Panel, Massey gives the following formal characterization of what he takes to be Riemann's conception of a singly extended manifold, a doubly extended and an n -ply extended manifold ([50], p. 332): a singly extended manifold is

"an ordered pair $\langle P, R \rangle$ consisting of a nonempty set P of at least two points and a serial order relation R on P . According as R is discrete, denumerable dense, or linear continuous . . . the manifold is itself said to be *discrete*, *denumerable dense*, or (*linear*) *continuous*. By a *doubly extended manifold* . . . , Riemann seems to have meant the Cartesian product $M \times M$ of a singly extended manifold M with itself.⁸ In general, by an *n*-ply extended manifold Riemann may be taken to have meant the n th Cartesian product of a singly extended manifold with itself."

Massey ([50], p. 343) then defines "an *interval of M*" by means of the serial relation R as "a subset I of P such that $(x)(y)(z) [(x \in I) \cdot (z \in I) \cdot Rxy \cdot Ryz \supset (y \in I)]$."

Massey's formal characterization of n -ply extended manifolds was not intended as a full-fledged explication of Riemann's conception of them. But even as far as it goes, I find myself both unwilling and unable to adopt Massey's formal characterization of an n -ply extended manifold for the following reasons:

1. Massey uses serial relations to characterize singly and n -ply extended manifolds. In my view, the single or multiple extendedness of any manifold which is a concrete realization of one of the spaces is an *intrinsic* feature of that manifold. Thus, the single extendedness of a linear continuous P -space would be an intrinsic feature of that particular physical space on my conception. But it will turn out in section 4 that, in general, *serial* relations on singly extended discrete P -space, on denumerably dense P -space, and on linear continuous P -space are *extrinsic* to these manifolds by my criterion, except for the special cases of P -spaces that have

⁸ For a brief discussion of the Cartesian product, see [8], p. 17.

exactly one extreme element. (Indeed, serial relations are always extrinsic by Massey's own *formal* criterion of extrinsicness ([50], p. 343, D3), although it differs in several crucial respects from my *material* criterion. One difference is that Massey's criterion, being formal, applies alike to abstract manifolds and to their [physical] realizations, whereas my material criterion does not pertain to abstract manifolds *as such* but to their "concrete" realizations.) Thus, from my point of view, it would be desirable to have a characterization of singly (and n -ply) extended P -manifolds, for example, that eschews serial relations. And since Massey's formulation does employ such relations for this purpose, I am unwilling to avail myself of it.

2. The points on the perimeter of any topological Euclidean circle form a space which is *one-dimensional* but is *not* a spatial system of serial order. This topological circle therefore fails to qualify as a singly extended manifold in Massey's sense, although a P -model or T -model of it does so qualify intuitively, I believe. I have thought all along that the topological circle should be included under the concept of a singly extended manifold for the following important reason: physical *time* in a suitable coordinate system or time-like geodesics might globally have the topology of the one-dimensional Euclidean circle and yet be thought of as singly extended. Hence I cannot accept any characterization of singly extended manifolds which excludes the topological circle, as Massey's does. Incidentally, Riemann himself would presumably have regarded this circle as a singly extended manifold. Consequently, Massey's characterization of n -ply extended manifolds does not lend itself to incorporation in a full-fledged explication of Riemann's conception of them.

Fortunately, my purposes allow me to forego the ambitious attempt to give a general characterization of a singly or multiply extended space by devices which would qualify as intrinsic to each physical realization of such a space in the sense of sections 2 and 4 below. And I can confine myself here to offering intrinsic renditions of the structures of the three particular singly extended P -spaces that will concern us in the sequel, viz. discrete P -space, denumerably dense P -space and linear continuous P -space. My account here of these three singly extended P -spaces will then permit me just to indicate how one would give an intrinsic characterization of the topological Euclidean P -circle, which is one-dimensional and was our fourth example of a singly extended P -space. But in section 2(c) my discussion of intrinsic *metrics* will include three-dimensional P -manifolds and the four-dimensional relativistic PT -manifold (space-time).

It will turn out that singly extended discrete physical space, denumerably dense physical space and linear continuous physical space each qualify *intrinsically* as a *system of betweenness* in the sense of the postulates given by E. V. Huntington ([39], pp. 2–3). Hence I shall use this triadic relation of betweenness in the characterization of these three P -spaces which I am about to give. And this procedure will enable me to eschew *serial* relations, since these dyadic relations are *extrinsic* to these P -spaces, as we shall see in section 4, except for the special kinds of P -spaces that have exactly one extreme element. Whether or not a serial relation is or is not intrinsic to, say, linear continuous T -space, i.e. to a certain kind of time, will depend upon whether there are suitable irreversible types of processes or not, unless this T -space has exactly one extreme element. (This point will be taken up in **Part B**.)

By contrast, serial relations will turn out in section 4 to be intrinsic to at least two of the three particular singly extended *I*-manifolds of arithmetic constituted by the natural numbers, the rational numbers and the real numbers.

Turning to our three singly extended spaces, we can give the following characterizations abstractly even though the criteria for the intrinsicness of the betweenness relation in any *P*-realization or *T*-realization are material ones:

(i) Discrete Space: A system of betweenness in the sense of Huntington [39] whose elements are metrically homogeneous and which satisfies the further condition that the cardinality of the set of elements between any two given elements is *finite*. Symbolically, if the ternary predicate *B* is used to denote the betweenness relation, " $(\hat{z}) \dots$ " denotes the set of *z*'s such that $\dots$, and " $\bar{\alpha}$ " denotes the cardinal number of the set α , then the further condition reads

$$(x)(y)[\overline{(\hat{z})B(x, z, y)} < \aleph_0].$$

I shall refer collectively to the conjunction of Huntington's betweenness postulates with this further condition as "the characteristic postulates." And I am indebted to J. A. Coffa for proving a result essential to this formulation of the properties of a discrete space: If a serial ordering obtains in any system of objects satisfying the above characteristic postulates, then the system in question will also satisfy the standard postulates for a discrete series given by Veblen and Huntington.⁹ In short, any model of our characteristic postulates will also be a model of Veblen's and Huntington's standard postulates for a discrete series, provided that a serial ordering is also present in the model. We shall see in section 4 below that if our discrete space is a *P*-space, then a serial ordering will need to be introduced *extrinsically*, except in the special case of a discrete *P*-space having exactly one extreme element, which will be discussed in the present Section (a).¹⁰

Several points should be noted concerning discrete space as defined. Firstly, the second of our defining conditions of a discrete space, i.e. the requirement that the cardinality of the set of elements between any two given elements be finite, is a *theorem* of Huntington's postulates for a discrete series ([40], section 23, p. 20). Secondly, the cardinality of a discrete space is either finite or *denumerably* infinite ([40], sections 24–27, pp. 21–23). Hence, our Huntingtonian kind of discrete space must *not* be equated with the so-called discrete *topological* space ([46], p. 66), *which allows other cardinalities*. Thirdly, if we define an "*extreme*" point of a discrete space as a point that is *not* between any two other points, then the discrete spaces fall into three species: spaces with two extreme points, spaces with one extreme point, and spaces without any extreme points. And it then follows from the definition of a discrete space that a discrete space with two extreme points is characterized by the fact that it is a *finite* discrete space. In other words, every discrete space with two extreme points is finite (by the second of our defining

⁹ These postulates for a discrete series, which are stated by means of the serial relation term "*precedes*" (or "*follows*") can be found, for example, in [40], p. 19.

¹⁰ For a statement of how a serial ordering may be introduced into a system of betweenness, see the Section "Serial order defined in terms of betweenness, with respect to *U*, *V*" in [39], p. 7.

conditions), and conversely every finite discrete space has two extreme points ([40], section 27, p. 23).

(ii) Dense Denumerable Space: A system of betweenness whose elements are metrically homogeneous and which satisfies the two further conditions that its subsets be countable and that it be dense (i.e. that *between* every two of its elements, there be at least one other).

(iii) Linear Continuous Space: The term "linear" here is intended to refer to the Euclidean *topological* straight line. To give our characterization of this space, we note first that Tarski ([67], p. 18) has given an axiom which renders the continuity of this space in the following remarkable way, beautifully suited to our needs: The axiom contains a betweenness predicate β as the only nonlogical constant, in addition to logical constants as well as first-order and second-order variables ranging over points and point-sets respectively.¹¹

Note that Tarski's continuity axiom does *not* involve any term denoting a serial relation; instead its only nonlogical constant denotes the betweenness relation. Hence we can say that linear continuous space is any system K of betweenness whose elements are metrically homogeneous and which also satisfies the following three postulates, *each* of which contains betweenness as the only nonlogical constant: the density postulate, Tarski's continuity axiom, and the postulate of linearity ([40], p. 44), which states that K contains a denumerable subclass R such that *between* any two elements of K there is an element of R .

Having used the betweenness relation to characterize three of our singly extended spaces, we can also use it to state what we mean by a closed interval in each of these spaces: the closed interval (a, b) contains the points a and b and all points *between* a and b , if any. In dense denumerable and linear continuous space, the open interval corresponding to the closed interval (a, b) is then obtained from the latter by excluding both end points a and b . And the two corresponding half-open intervals are obtained by excluding exactly one of the two end-points at a time from the closed interval. In a *multiply* extended space, the specification of a singly extended interval (a, b) could be given by reference to a suitable singly extended subspace of the multiply extended space. For example, an interval (a, b) linking two points a and b in three-dimensional physical space, or two punctal events in four-dimensional space-time, could be specified by reference to one of the particular one-dimensional, linear proper subspaces (or linear continuous paths) connecting a and b . Such a singly extended interval in a multiply extended space must, of course, be distinguished from the multiply extended kind of interval in the same space, since the latter type of interval is itself multiply extended. The so-called "generalized rectangle" ([42], p. 200) is an example of a multiply extended interval.

¹¹ It is worth noting that O. Veblen [69] gives a presentation of Euclidean geometry in which one of the nonlogical primitives is a ternary "betweenness" predicate " $\{ABC\}$." But whereas Veblen's relation $\{ABC\}$ and Huntington's betweenness relation ABC hold only when A , B , and C are pairwise distinct, Tarski's relation $\beta(x, y, z)$ holds when y is identical with x or z ([67], p. 17).

Accordingly, when we combine Huntington's postulates for betweenness with Tarski's continuity axiom, we assume that we have defined Tarski's betweenness in terms of Huntington's.

Though the singly extended and multiply extended senses of "interval" are clearly different, the term "interval" refers to a set of punctal elements in both of these two uses. Hence these two uses must be distinguished from a third use in which the term "interval" denotes the values of a *numerical* function on ordered pairs of punctal elements of a space. Thus, when discussing the Minkowski space-time of special relativity, Landau and Lifshitz ([44], Ch. 1, Section 1-2, pp. 4-9) devote a section of their book to "intervals" in this third sense. Referring to a Minkowskian coordinate system, they write ([44], p. 5): "If $x_1y_1z_1t_1$ and $x_2y_2z_2t_2$ are coordinates of *any* two events, then the quantity

$$s_{12} = [c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2]^{1/2}$$

is called the *interval* between these two events." Note that in this third usage, the term "interval" denotes a *number*, and *not* the pair of points to which it is assigned, let alone a linear continuum of points connecting these two points (and containing them as members)! And when Landau and Lifshitz call the numerical values s_{12} "the interval between" their two punctal events a and b , the term "between" *does not* denote a relation among punctal events; instead this usage of "between" merely conveys the following *symmetry*: the numerical value s_{12} is assigned alike to each of the two ordered pairs of events a, b and b, a . "Between" can be used to convey symmetry, because betweenness is governed by the symmetry condition $ABC = CBA$ ([39], Section 2, p. 2).

One reason for distinguishing the three different senses of "interval" above is that writers on relativity theory, such as A. S. Eddington, have traded on the third *numerical* sense of "interval" to suggest misleadingly that they have specified a linear continuous set of events which connect two given punctal events a, b in 4-dimensional space-time. Thus, Eddington first generalizes Landau and Lifshitz's *numerical* meaning of "interval" to an arbitrary space-time by using the quadratic differential form $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ to define "the interval ds between two neighboring events with coordinates (x_1, x_2, x_3, x_4) and $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3, x_4 + dx_4)$ in any coordinate system" ([18] p. 10). But after giving this *numerical* definition, he switches to a *geometrical* meaning of "interval" when he declares ([18], p. 11): "The quantity ds is a measure of the interval. It is necessary to consider carefully how measure-numbers are to be affixed to the different intervals occurring in nature." And immediately after again speaking numerically of "the interval ds between two points" ([18], section 93, p. 216), Eddington goes on to appeal to a tacit geometrical meaning of this expression by writing ([18], section 93, p. 217): "The adoption of a particular tensor $g_{\mu\nu}$ is equivalent to assigning a particular gauge-system—a system by which a unique measure is assigned to the interval between every two points" (cf. also [18], p. 200). In the same vein, the mathematician C. E. Weatherburn speaks of "the linear element ds ," and hence appears to assign both a numerical and a geometrical meaning to " ds ," when he says "Thus the square of the linear element ds is given by a quadratic form in the differentials of the coordinates" ([70], p. 35). Landau and Lifshitz ([44], p. 6) apply the term "line element" to the *number* given by the product of $\sqrt{-1}$ and their particular ds .

What of the structure of the one-dimensional topological Euclidean circle? A characterization of this space would presumably contain, among others, the statement that this space is a system of "separation of pairs" of points in the sense of Huntington ([39], Section 4, pp. 4–5). For in the interesting case of "circular" (closed) time (*T*-space), the tetradic relation of separation turns out to qualify as intrinsic in the sense of section 2 and section 4 below. But the separation postulate is insufficient to characterize the full-blown structure of the one-dimensional Euclidean topological circle. To supply a characterization suited to our needs when it is applied to a possible temporal model, further conditions involving only intrinsic non-logical devices would have to be added. Recall our addition above of three requirements to Huntington's betweenness postulates to give a description of linear continuous space which would qualify as intrinsic in a *P*-realization. Given the appropriate kind of rendition of the properties of the Euclidean topological circle, it is then possible to characterize the (two) intervals or "segments" (*a*, *b*) that connect a pair of fixed points *a* and *b*, as is done by Heyting ([37], pp. 114–115).

I have been discussing here in section 1(a) characterizations of the *intervals* of our four singly extended spaces by means of relations that will turn out to qualify as intrinsic to particular physical realizations of these spaces. For this discussion has set the stage for our consideration, in section 2(c) below, of the *criteria* for the intrinsicness or extrinsicness of a *metric* to the intervals of a concrete manifold and to the concrete manifold as a whole.

When speaking above of singly extended spaces which are respectively discrete, dense denumerable, and linear continuous, I remarked that section 4 below will show spatial serial relations to be extrinsic to *P*-models of these three spaces, unless they have exactly one extreme point. Let me now call attention to the important respects in which those species of these three spaces that do have exactly one extreme element are remarkable even abstractly. The possession of exactly one extreme element is noteworthy in the case of *any one of our three spaces* for the following reasons: (i) it permits us to use Huntington's betweenness relation as the only nonlogical relation to identify uniquely (individuate) the one extreme point E_1 of the space, and (ii) as Massey has pointed out to me, Huntington's betweenness can then be used furthermore to define a serial order by reference to the extreme point E_1 .

Thus, the extreme ("first") point E_1 can be defined as the unique point which is not between any two others:

$$E_1 = (\iota x) \sim (\exists y)(\exists z)B(yxz),$$

where *B* denotes Huntington's betweenness relation. For example, consider a linear continuous space that has exactly one extreme element. As I have noted elsewhere by reference to a *temporal* model of that space ([26] and [27], p. 13, fn. 7), no ostensive device is needed to individuate the singular instant of the big bang in a big bang model of cosmology that has an infinite future. For the one extreme instant E_1 of that model can be characterized by means of betweenness, as we have just seen. Given this identification of the point E_1 by means of betweenness, Massey then defines a serial relation $P(x, y)$ of the points of the space as follows:

$$P(x, y) =_{\text{def.}} \{(x = E_1) \cdot (y \neq E_1)\} \vee B(E_1xy).$$

Let us call any one of our three singly extended spaces that has exactly one extreme element an "*E*-space." Then, as section 4 will explain, if the betweenness relation qualifies as intrinsic to a particular model of one of the *E*-spaces, then so also does the corresponding serial relation $P(x, y)$.

Note that no such definition of a serial relation is feasible in the case of those singly extended discrete, dense denumerable and linear continuous spaces that have two extreme elements, which I shall call "*F*-spaces." For in the case of an *F*-space, neither of the two extreme points is between any two other points. And since *each* of these two extreme points is characterized by the property of failing to be between any two other points, the uniqueness presupposition of the definite description E_1 fails to obtain. Hence in the case of an *F*-space, Massey's definition of $P(x, y)$ defines a predicate having an empty extension.

Suppose, however, that we disregard the devices by which elements are identified and are given two distinct, identified elements in *any* system of betweenness, be it an *F*-space or one of our three singly extended spaces without any extreme points. In the case of two given, identified elements, Huntington's or Tarski's betweenness relation can indeed be used to specify a serial ordering of all the elements of the system of betweenness. Huntington ([39], p. 7, section 3.1), Tarski ([67], pp. 21-22), and Massey ([50], p. 339) have given us three different ways in which this specification can be given.

It should also be pointed out that within the class of *E*-spaces, singly extended discrete space with exactly one extreme element is distinguished from the dense *E*-spaces by the following fact: Although its elements are (normally) held to be homogeneous in regard to metrical extension, the Huntingtonian betweenness relation on that space suffices to individuate* *all* of its points rather than merely the one extreme point. To see this, first recall how we uniquely identified the one extreme ("first") point E_1 by means of betweenness in any one of the *E*-spaces. The next ("second") point E_2 is also characterizable in terms of betweenness:

$$E_2 = (ix) [\sim (\exists y) B(E_1 y x) \cdot x \neq E_1]$$

Then the $(k + 1)^{\text{st}}$ point *next* to the k^{th} point can be identified for $k > 1$ by a definition given by Massey:

$$E_{k+1} =_{\text{def.}} \text{the element such that } E_k \text{ lies between it and the first element } E_1, \text{ and no element lies between it and } E_k.$$

Thus, any model of singly extended discrete space with exactly one extreme element is remarkable for the fact that *all* of its points can be individuated by means of the betweenness relation. Therefore, models of singly extended discrete *E*-space qualify as inhomogeneous manifolds ("*I*-manifolds") with respect to betweenness, while being homogeneous with respect to metrical extension. In particular, the purely geometric relations of a *P*-model of a discrete *E*-space suffice to *individuate all* of its point elements even though these elements are *homogeneous* with respect to metrical extension. By contrast, it is clear that for an enormous range of other kinds of *P*-spaces, the purely *geometric* properties and relations of these spaces will

* As noted on p. 465, "individuate" in the sense of identifying uniquely or picking out.

not suffice to individuate *all* of their metrically homogeneous points. And hence these points will depend for their individuation on (monadic and/or polyadic) *physical attributes which are extra-geometric*. This conclusion seems to me to pose a fundamental philosophical question for the Clifford-Wheeler vision of a “space-theory of matter” which J. A. Wheeler has called “*geometrodynamics*.”

See
Append.
§ 46

For the hope of this vision is “to understand matter as a manifestation of empty, curved space” or as “*an excitation-state of a dynamic geometry*,” leading to a “purely geometric description of nature” ([74], p. 1; the translation from the German text is mine). On this view, “There is nothing in the world except empty curved space. Matter, charge, electromagnetism, and other fields are only manifestations of the bending of space. *Physics is geometry*” ([73], p. 225). Such a description is clearly more ambitious than the original formulation of general relativity in which the 4-dimensional geometry prescribes the space-time paths of matter-particles and light rays. For in that original formulation, autonomous matter, in turn, at least affects the curvature of space-time even if it does not wholly determine it without additional boundary conditions. Thus, in the original formulation, matter (or energy) was conceived as determining the temporal evolution of the spatial geometry in a given 4-dimensional coordinate system, if the spatial geometry is given at some initial instant. And entities foreign to the metric geometry of space-time were held to play an autonomous role, since particles and fields *other than* gravitation were regarded as precisely such entities. In this sense, the original conception of general relativity was more that of a matter theory of space than of a space theory of matter.

Hence the basic philosophical question for the “space-theory of matter” in the Clifford-begotten Wheeler vision of geometrodynamics is: What is its principle of individuation? For, if the *extra-geometric* physical entities and attributes of *prior* conceptions are to be understood as *intra-geometrodynamical* ones, then what serves to individuate the metrically homogeneous punctal event elements of the space-time manifold? In other words, it would seem that a space-theory of matter *cannot* avail itself of the individuating devices of the *usual* conception of relativity in which a space-time point was held to be individuated by the intersection of the world lines of two test particles or of one test particle and a light cone, or the like.

(b) *Manifolds whose elements are intrinsically inhomogeneous with one another: I-manifolds with respect to a set C of intrinsic properties*

We remarked earlier that certain serial relations (i.e. “less than” and “greater than”) will turn out to be intrinsic to at least two of the three particular singly extended manifolds of arithmetic constituted by the natural numbers, the rational numbers and the real numbers. The intrinsicness of these serial relations will be seen to be rooted in the fact that the number elements of these manifolds, qua arithmetic objects, are inhomogeneous with one another with respect to familiar arithmetic properties that are intrinsic to them.

No wonder therefore that intrinsic serial relations of less than and greater than, and thereby also an intrinsic relation of betweenness, are each available in the case

of arithmetic to define *intervals* in each of the three aforementioned arithmetic manifolds. Thus, for example, the closed interval (a, b) in the manifold of real numbers is the set of all reals x such that $a \leq x \leq b$ or $a \geq x \geq b$. And the intrinsic serial relation $<$ among the reals can then be used to define Huntington's betweenness relation ABC in this number manifold as follows ([39], p. 6):

$$ABC =_{\text{Def.}} \{(A < B) \cdot (B < C)\} \vee \{(C < B) \cdot (B < A)\}.$$

Clearly, this numerical betweenness qualifies as intrinsic, since the numerical relation $<$ does so qualify, as will be seen in section 4.

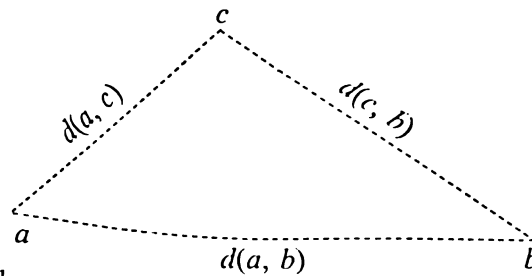
2. Intrinsicness vs. Extrinsicness of Metrics, Metrical Equalities, and Congruences

(a) *Distance metrics, measure-metrics, and metrics of a Riemannian space*

It behooves us to distinguish between different kinds of metrics as follows: Distance-metrics, which I shall call "*D-metrics*," measure-metrics, to which I shall refer as "*M-metrics*," and metrics of a Riemannian space, which I shall designate as "*R-metrics*." Correlatively, we shall need to distinguish between several kinds of metrical equality, and to point out that the term "congruent" is *not* used for denoting *every* kind of metrical equality of sets. Furthermore, we shall have to be mindful of the fact that the usage of the terms "distance," "length" and "metric" is by no means uniform in the mathematical literature. For example, some authors reserve the term "metric" for only those metrics that we are about to define as "distance metrics." And they then reserve the term "metric space" for a space on which a distance function is defined.¹² I apply the term "metric" alike to distance functions, measure functions and to the "length" functions of a Riemannian space. In so doing, I am using the term to refer to at least the following three kinds of *gauge functions* on specified types of sets:

1. Distance-metrics ("*D-metrics*"). Let X be a nonempty set. A real-valued function d defined on ordered pairs of elements in X is called a "*distance function*," "*distance-metric*" or—for brevity—"D-metric" on X , iff it satisfies the following axioms for every $a, b, c \in X$:

- (i) $d(a, b) \geq 0$ and $d(a, a) = 0$.
- (ii) (Symmetry) $d(a, b) = d(b, a)$.
- (iii) (Triangle Inequality) $d(a, b) + d(b, c) \geq d(a, c)$.



- (iv) If $a \neq b$, then $d(a, b) > 0$.

The real number $d(a, b)$ is called the "*distance*" from a to b .

¹² Cf. the definitions of "metric" and of "metric space" in [42], p. 235, and in [46], p. 111.

We shall need to emphasize the similarities and differences between these axioms and those to be given below for measure-metrics and Riemann-metrics. Note therefore that axiom (i) asserts that the distance from any point to another can never be negative *and* that the distance from a point to itself is zero, while axiom (iv) states that the distance between two distinct points is positive. The symmetry asserted by axiom (ii) justifies the locution "the distance *between* a and b ." But we must guard against allowing this use of the term "between" to suggest misleadingly that the entities on which distance functions themselves are defined generally qualify as singly extended intervals. The reason for the designation "triangle inequality" for axiom (iii) is obvious from the application of this axiom to a spatial situation of the kind depicted in the diagram adjoining its statement above.

With respect to a particular D -metric, set A is said to be *congruent* to set B iff there is a 1-1 mapping ϕ of A onto B such that for any $a, b \in A$, $d(a, b) = d[\phi(a), \phi(b)]$. Since congruence is a symmetric relation, we shall speak of sets A and B as being congruent to one another. Transformations of the kind ϕ mentioned in the definiens of "congruent" sets are called "congruent transformations" ([38], p. 166). Since distances are defined on pairs of points, a D -metric generates a *quaternary* relation of equidistance, i.e. a relation a, b, c, d such that a is as distant from b as c is from d . The relation of congruence among sets is defined with respect to a D -metric, which in turn is defined on pairs of points. But the sets which qualify as congruent do not, of course, have to be sets containing two elements but can also be singly extended or doubly extended *intervals* in denumerable dense or continuous space, for example.

2. Measure-metrics (" M -metrics"). Let Y be a collection of subsets of a set X which constitute a ring of sets ([42], p. 314). Let M be a single-valued function defined on Y . Then ([42], pp. 232-233):

(a) The set function M is a *finitely additive measure* iff

- (i) for each member y_i of Y , $M(y_i)$ is either a non-negative real number or $+\infty$
- (ii) $M(\Lambda) = 0$
- (iii) if the sets $\{y_i\}$ in Y are *disjoint* for different values $i = 1, 2, \dots, n$, then $M(\bigcup_{i=1}^n y_i) = \sum_{i=1}^n M(y_i)$, where it is to be understood that any sum is equal to $+\infty$ if one of the summands is $+\infty$.

(b) The set function M is a *countably additive measure*, iff M is a finitely additive measure *and*

- (iv) if the sets $\{y_i\}$ in Y are disjoint for different values $i = 1, 2, \dots$, and if the union $\bigcup_{i=1}^{\infty} y_i$ of all sets $\{y_i\}$ belongs to Y , then $M(\bigcup_{i=1}^{\infty} y_i) = \sum_{i=1}^{\infty} M(y_i)$, where it is to be understood that if $\sum_{i=1}^k M(y_i)$ is not bounded as a function of k , then $\sum_{i=1}^{\infty} M(y_i) = +\infty$, and also that any sum is equal to $+\infty$ if one of the summands is $+\infty$.

Strictly speaking, the distance function assigns the distance zero *not* to a singleton, but to the ordered pair (a, a) , i.e. to the distance of any element from itself. On the other hand, measure functions assign numbers to singletons as well as to sets of infinitely many other cardinalities. Despite the logical distinction between an ordered pair (a, a) and a singleton $\{a\}$, there are applications of D -metrics and M -metrics in which the distance of a point from itself and the measure of a singleton play corresponding roles. In regard to such a context, note that whereas axiom (i) for the D -metrics requires that the distance of a point from itself be zero, the corresponding axiom for M -metrics requires only that the M -gauge of a singleton be non-negative. And while the distance axioms (i) and (iv) confine the value zero to the distance of an element from itself, a measure function M neither is required to assign zero to a singleton, nor is M required to confine the assignment of zero to a singleton, if it does assign it. Furthermore, whereas a D -metric on X is defined on ordered *pairs* of elements in X , the subsets $\{y_i\}$ of X on which M -metrics are defined *are in no way* restricted to sets containing two elements. But at least in the familiar context of plane Euclidean geometry, the following restriction governs any M -metric on a set X on which a D -metric is already defined: If two sets qualify as *congruent* with respect to the D -metric in the sense defined above, then the M -metric must assign the same measures to them ([38], p. 166). Yet sets which are assigned the same measures by the M -metric need *not* be congruent. The case of rectangles in the Euclidean plane will illustrate these statements.

Suppose that in the Euclidean plane, we are given the standard Euclidean distance-metric corresponding to a particular choice of unit and the corresponding length function on the linear intervals. Then the standard Euclidean area-numbers assigned to rectangles in this plane qualify as values of a permissible measure-metric ([38], p. 165 and [42], p. 18). Two 6 by 2 rectangles, for example, are congruent and also have equal area-measures of 12. But a 3 by 4 rectangle and a 6 by 2 rectangle are *incongruent*, although they have the same area-measures 12. Thus, here there is *metrical equality* with respect to a given familiar M -metric (area), although there is *no* congruence or metrical equality with respect to a specified familiar D -metric.

To convey this discrepancy in equality with respect to the specified D -metric and M -metric, I shall say *for brevity* that the two rectangles are " M -equal" but *not* " D -equal." By the same token, there are *solid* figures in Euclidean 3-space that are incongruent but have equal volume-measures. And thus our examples here illustrate that the predicate "congruent", when applied to sets, is generally *not* synonymous with the predicate "metrically equal": the application of the term "congruent" is relativized to a particular D -metric, whereas the use of the term "metrically equal" is relativized to instances of other kinds of metrics as well.

Indeed, there is at least one permissible (albeit trivial) *measure* function M_0 on the subsets of singly extended discrete denumerable space with respect to which *all* these subsets are M -equal, although there is *no* allowable *distance* function *whatever* with respect to which these same subsets will all be congruent to one another. For, all subsets $\{y_i\}$ are M -equal with respect to the *constant* measure function $M_0(y_i) = 0$. But since the 1-1 mappings required for congruent transformations do not exist

in the case of subsets of different cardinality, there is no D -metric whatever with respect to which all the M_0 -equal subsets are congruent.

On the other hand, we shall soon see that the infinitesimal numbers ds furnished by the so-called positive definite R -metrics qualify as *lengths* of *linear intervals* because they qualify as distances between their end-points. Thus, positive definite R -metrics ds furnish the values of distance functions and of length functions. And two *infinitesimal* intervals which have the *same lengths* with respect to a particular positive definite R -metric ds will also be *congruent* with respect to the D -metric furnished by that function ds .

Comparing the triangle inequality axiom governing D -metrics with axiom (iii) of finite additivity or axiom (iv) of countable additivity for M -metrics, we note an important difference: The ordered pairs (a, b) and (b, c) whose distance numbers $d(a, b)$ and $d(b, c)$ are added in the triangle inequality are *not* disjoint but have the common element b ; by contrast, the additivity axioms (iii) and (iv) for the M -metrics make the additivity of the measures of two or more sets conditional on their being disjoint.

3. Riemann-metrics ("R-metrics"). In a space of n dimensions, coordinatized by ordered n -tuples of real numbers, consider any *two* points whose coordinates in any such coordinate system are x^i and $x^i + dx^i$ ($i = 1, 2, \dots, n$). Let us employ the Einstein summation convention according to which the appearance of the same index in any term as both a subscript and a superscript signifies that this term stands for the sum of all the terms obtained by giving that index all the values it may take ([70], p. 2). Then we can say that Riemann defined his kind of metric on a coordinatized n -space by means of a quadratic form in the differentials of the coordinates by the equation

$$ds^2 = g_{ij} dx^i dx^j \quad (i, j = 1, 2, \dots, n),$$

where the coefficients g_{ij} are functions of the coordinates x^i subject only to the restriction of having a nonvanishing determinant. Thus, the Riemannian metric ("R-metric") ds is defined on *pairs* of points, although in the case of a space of two or more dimensions, various authors (e.g. Eddington) contrive ds as being assigned to a particular linear continuous path linking the pair of points, i.e. to an *interval* or *arc* of which the pair of points are the termini. And they then invoke *this* construal, as Reichenbach does ([58], pp. 178, 180), to justify speaking of ds as a "length," albeit in a *generalized* sense of that term.¹³

Note at once that as defined by Riemann, there is *no* requirement that the values of his metric functions $ds = \sqrt{g_{ij} dx^i dx^j}$ be *real* numbers for arbitrary values of the coordinate differentials, since ds may be imaginary. By contrast, the values of both the distance functions and the measure-functions were confined to the non-negative *real* numbers.

¹³ Thus Reichenbach emphasizes in this connection that ds and, on geodesics, the non-infinitesimal Δs , are assigned to *intervals* rather than to their end-points by writing: "By the length of a line-segment we do not understand the segment itself, but a number coordinated to it. A segment is determined if we give the coordinates of the *totality of points which lie on it*" ([58], p. 178, my italics).

Thus, consider the familiar special case of the 4-dimensional Riemann space constituted by the space-time of Einstein's special relativity. In a Minkowskian coordinate system, a particular R -metric of that space-time is given by

$$ds = \sqrt{dx^2 + dy^2 + dz^2 - c^2 dt^2},$$

and an *alternative* one is specified by

$$d\tau = \sqrt{dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)},$$

where we may let the speed c of light be unity. These R -metrics of Minkowski space-time have the important feature that their metric coefficients g_{ij} do *not* all have the same sign: For a choice of units such that $c = 1$, each of these metrics has at least one coefficient whose value is $+1$ and at least one other coefficient whose value is -1 . Since the signs of its coefficients are *not* all the same, each of these metrics is said to be "indefinite" ([70], p. 12). If, *contrary* to fact, all of the signs had been the same in either metric, then it would have been called "definite." And in the latter case, it could have been either "positive definite" or "negative definite," depending on whether the signs had all been positive or all been negative respectively ([21], p. 23). But the numbers ds or $d\tau$ assigned to pairs of elements by an R -metric will be real for *arbitrary* values of the coordinate differentials only in the case of a positive definite metric ([21], p. 35).

Hence for each of our two particular space-time R -metrics ds and $d\tau$, no less than for any of the other permissible R -metrics on Minkowski space-time, there will be certain pairs of events to which it assigns *imaginary* numbers ([44], p. 7). And in the case of our R -metric $d\tau$, for example, the familiar distinction between a pair of events whose separation is said to be "space-like" and another pair whose separation is said to be "time-like" is correlated respectively with whether $d\tau$ is imaginary or a positive real number.

Interestingly enough, authoritative treatises by mathematicians and physicists refer to the numbers assigned to pairs of elements by R -metrics (including the indefinite ones of relativity) as "distances" ([44], p. 5, and [70], p. 35), and as "elements of length" ([21], p. 34) or "elements of arc length ds " ([42], p. 312). This use of the terms "distance" and "length" is noteworthy not only because the functions allowed by R -metrics may have imaginary values but for a number of important additional reasons:

(1) Consider two distinct events in Minkowski space-time whose 4 coordinate differences are infinitesimals dx^i ($i = 1, 2, 3, 4$) such that their space-time separation

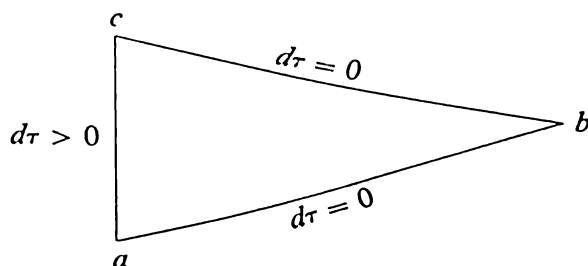
$$d\tau = \sqrt{dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)}$$

is *zero*. Pairs of events having the *null* separation $d\tau = 0$ are, of course, connectible by a direct (unreflected) light ray in vacuo. Consider next any two events whose 4 coordinate differences Δx^i ($i = 1, 2, 3, 4$) are *not* infinitesimals and which are connected by a direct light ray in vacuo. The latter light ray is a 4-dimensional null

geodesic along which $d\tau = 0$ everywhere for any pair of events having only infinitesimal coordinate differences. Hence the space-time separation τ of any pair of events having *noninfinitesimal* coordinate differences and connectible by a direct light ray in vacuo is *zero*. For in the case of any such pair, τ is given by $\int d\tau = 0$, the integration being carried out along the null geodesic. Thus, we see that *any* two events whatever that are connectible by a direct light ray in vacuo have a space-time separation of zero.

This result shows that there are pairs of distinct elements a, b in Minkowski space-time whose 4-dimensional separation violates the requirement of axiom (iv) for D -metrics that $d(a, b) > 0$ for $a \neq b$. (Indeed, this violation obtains in curved space-times of general relativity no less than in the flat space-time of Minkowski). To be sure, the above indefinite R -metrics ds and $d\tau$ assign the value zero as the separation of an event from itself, as required by $d(a, a) = 0$. But these metrics *also* assign zero to any pair of distinct events belonging to a light ray. By contrast, the axioms for D -metrics require that zero be assigned only as the distance of a point from itself.

(2) Moreover, the permissible indefinite R -metric $d\tau$ violates the triangle inequality stated in axiom (iii) for D -metrics. To see this, consider the emission a of a light ray at a certain time from a space point A of an inertial system, its reflection b by a mirror at another space point of that system, and the ray's return c to the original space point A at a certain later time. The diagram shows the world-line $\overline{ac}$ of the stationary clock at A , while the lines $\overline{ab}$ and $\overline{bc}$ represent the world-lines of the outgoing and returning light-rays respectively. Clearly $d\tau(a, c) > 0$, since $d\tau(a, c) = dt$, and $dt > 0$. But $d\tau(a, b) = d\tau(b, c) = 0$. Hence the metric $d\tau$ violates the triangle inequality for D -metrics.



Books on Riemannian geometry explicitly speak of Riemann's ds as the "metric" of his n -dimensional space ([21], p. 34, and [70], p. 35). But writers on "metric spaces" often restrict that term to spaces on which a *distance* function satisfying the axioms for a D -metric above has been defined ([42], p. 235). The latter usage would prevent a Riemannian space from qualifying as a metric space.

(3) In certain familiar physical contexts, the length l of a closed linear *interval* which is *infinitesimal* defines the distance between its end-points as follows: If I

is the set of points of the closed linear interval $[x, y]$, then $d(x, y) = l(I)$ in these contexts. Here the length l is obtained by applying the rod to the *infinitesimal interval* $[x, y]$. And the number l thus obtained is then also assigned to the termini of the interval as the distance $d(x, y)$ between them. For the infinitesimal interval coincides with (part of) the rod and is thus a "direct" path, or path of least *length* linking its end-points. The length ds of an infinitesimal interval, thus measured, can then be expressed by a positive definite R -metric. But the fact that the ds function depends only on the differences between the coordinates of the two *end-points* and (generally) also on the coordinates of one of them cannot detract from the following other fact: the coincidence between an ordered pair of points A, B of the space with two points of the rod—which renders the distance from A to B physically determinate—involves the coincidence of (part of) the rod with an *interval* $[A, B]$ of which A and B are the termini. Thus, in a Riemann space whose R -metric is *positive definite*, the infinitesimal ds qualifies as both a length function and a distance function, if it is understood that the sign of the square root of the quadratic form is always taken to be positive. In the context of positive definite R -metrics ds , the length numbers ds also qualify as distances *even in a space of two or more dimensions*. For, as we saw, the closed linear intervals to which these numbers are assigned are infinitesimal and are therefore "direct" paths between their end-points. In a dense space, the length of a closed interval is also taken to be the length of the associated open and half-open intervals. Hence the numbers ds of positive definite R -metrics qualify as lengths of the latter kinds of infinitesimal intervals as well. And since the measures of intervals can be given by their lengths, a positive definite kind of R -metric ds furnishes the measures which an allowable M -metric assigns to intervals.

But Riemannian geometers use the term "element of *arc length* ds " to refer to the numbers assigned by *any* permissible R -metric, including the indefinite ones of relativity theory and negative definite R -metrics ([21], p. 34). And as noted by Reichenbach ([58], p. 178), this extended use rests on a consistent generalization of the concept of length to cover the latter cases as well.

Hence in this context, it is incorrect to say that length functions are instances of measure functions and accord with them. Since the Riemannian arc length can be imaginary in some cases, it is *not* true here that for a set which is an interval, the length of the set will qualify as a measure of the set in every case.

(b) *Criteria for the nontriviality of several species of metrics*

1. General remarks. I shall be discussing requirements for the nontriviality of each of our various metrics, beginning with M -metrics in the next subsection. Before doing so, I wish to note the distinction between the concept of a trivial metric—which is to be discussed below—and the concept of a *trivial difference* between metrics that may or may not themselves be trivial. And I shall briefly illustrate this distinction by reference to D -metrics. But the requirements for the nontriviality of D -metrics as such will first be discussed after we have considered both M -metrics and length functions.

For a set X of at least three elements, let d_1 be the distance function on X defined by

$$d_1(a, b) = \begin{cases} 0 & \text{if } a = b \\ 1 & \text{if } a \neq b. \end{cases}$$

And let d_2 be the different distance function on X defined by

$$d_2(a, b) = \begin{cases} 0 & \text{if } a = b \\ 2 & \text{if } a \neq b \end{cases}$$

Mathematicians consider each of these distance functions to be a *trivial* D -metric ([46], pp. 111 and 116), since it assigns the same number to all ordered pairs (a, b) of distinct points. Thus, suppose that a, b , and c are three distinct points of a linear continuous P -space (see section 1(a) above) such that b is between a and c . Then each of these two trivial D -metrics is insensitive to the fact that b stands in the relation of spatial betweenness to a and c . I shall argue later on that this spatial insensitivity is sufficient to render these D -metrics trivial.

Hence it will turn out that a minimal necessary condition for the nontriviality of a D -metric on such a set X is that it may *not* assign the same distance to all pairs of distinct points.

It is very important to note that two D -metrics can each be *nontrivial* while *differing* from one another only trivially. For example, in a Euclidean P -plane coordinatized by a fixed network of Cartesian coordinates x, y , consider the two Euclidean distance functions

$$D_1(a, b) = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}$$

and

$$D_2(a, b) = k \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2},$$

where k is some multiplicative positive real constant different from 1. Although neither of these D -metrics is itself trivial, they differ only trivially, since their difference is solely a matter of the choice of unit and does not affect the *ratios* of distances. Needless to say, there are also nontrivial D -metrics on, say, a two-dimensional coordinatized space which *differ nontrivially*. Thus the two nontrivial D -metrics

$$ds_1^2 = dx^2 + dy^2$$

and

$$ds_2^2 = dx^2 + x^2 dy^2$$

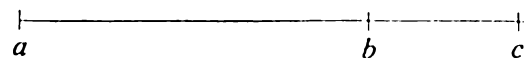
are each Euclidean ([24], pp. 99–100) and yet they differ nontrivially. But two trivial D -metrics such as our d_1 and d_2 above can differ only trivially.

2. M -metrics. In laying down necessary conditions for the nontriviality of M -metrics, we shall confine ourselves to restricting the values of those measures which are generalizations of length or correspond to length in singly or n -ply extended P -manifolds or in singly extended T -manifolds. Although it is not required for our

purposes, the extension of our necessary conditions for nontriviality to the case of measure functions whose values are generalizations of areas or volumes of P -space is, of course, no less desirable.

For the reasons explained at the end of section 2(a) above, the positive definite R -metrics, construed as being defined on linear intervals, can furnish permissible measures of them. We can therefore speak meaningfully of the values ds assigned by positive definite R -metrics as fulfilling the non-triviality conditions which we shall impose on the values of measure functions. But when we deal separately with *indefinite* R -metrics, we shall see that they satisfy only a *weakened* counterpart of the first of our two necessary conditions for the nontriviality of an M -metric.

As Riemann noted when discussing measurement in his Inaugural Dissertation, in a continuous space of points, comparison of, say, linear intervals is possible without a specific metric to *some* extent as follows: If one of two such intervals "is part of the other, . . . one can . . . decide upon the question of more and less," though "not upon the question of how much" ([60], p. 274). Thus, consider the sets of linear continuous intervals of a two- or three-dimensional P -space and of a linear one-dimensional T -space. In one of these sets, take two closed intervals (a, b) and (a, c) such that the former is a proper sub-interval of the latter as depicted here.



Clearly the spatial (or temporal) interval (a, c) includes (a, b) as a proper sub-interval such that their difference is itself a non-degenerate interval. Hence to say that (a, c) *extends over* (a, b) by a *half-open interval* is to say that all of the members of (a, b) also belong to (a, c) , but that (a, c) includes a (half-open) interval whose elements do not belong to (a, b) . Corresponding remarks apply to the case in which one closed interval (a, d) extends over another closed one (b, c) by two separate half-open intervals (a, b) and (c, d) . A measure function which does not assign a different number to a closed interval than to *every* one of its proper closed sub-intervals, including singletons, would be *insensitive to the stated facts of spatial (or temporal) extension or inclusion*. And by thus failing to tell the spatial or temporal story, as it were, such a measure function would be *trivial*.

For example, suppose that a countably additive measure function on one of our P - or T -spaces were to make the following assignments based on cardinality: Zero to a singleton, and hence also zero to finite and to denumerably infinite sets; $+\infty$ to any noncountably infinite point set. By assigning $+\infty$ or 0 respectively to all nondegenerate intervals alike according as the manifold is continuous or not, this cardinality-based M -metric would be trivial.

Let us now state a necessary condition for the nontriviality of M -metrics by reference to *closed* intervals, including singletons. This necessary condition will apply to the closed intervals of a singly extended *discrete* denumerable P - (or T -) manifold (cf. section 1(a) above), no less than to the closed intervals of dense denumerable, or continuous 3-dimensional space, for example.

If our universe of discourse is the set of all *closed* intervals of the particular space X , including singletons, then a necessary condition for the nontriviality of an M -metric on X is that

$$(x)(y)\{[(x \subset y) \cdot (x \neq y)] \supset M(x) \neq M(y)\}.$$

The fulfillment of the condition $M(x) \neq M(y)$ for the case of all closed intervals x, y such that $(x \subset y) \cdot (x \neq y)$ will automatically assure, via the axioms governing finitely additive measures, that $M(y) > M(x)$. Since any countably additive measure is also finitely additive, any results that we establish for finitely additive measures will also apply to countably additive measures. In the sequel, we shall *not* assume that our measure functions are countably additive, rather than merely finitely additive, unless this is stated explicitly. But we shall assume on physical (spatial or temporal) grounds that in any P - (or T -) manifold, *length* is countably additive in at least the following weak sense of Massey, which pertains to the lengths of intervals lying "end-to-end": Consider any interval I which is the union of a finite or denumerable sequence of subintervals such that each member of the sequence has exactly one point in common with its predecessor, if any, and also with its successor, if any, and furthermore is disjoint from all the other members of the sequence. Then the length $L(I)$ is the arithmetic sum of the lengths of these subintervals.¹⁴

Note that a measure function or a length function satisfying our necessary condition for nontriviality is sensitive to such betweenness relations as exist among the end-points of the closed intervals to which it pertains. Thus, consider all the points x *between* the end-points (a, b) of a closed interval on the Euclidean line. The fulfillment of our necessary condition assures that for all such points x , all closed intervals (a, x) and (x, b) are assigned smaller measures than (a, b) . And precisely because measure functions (and length functions) satisfying our necessary condition here are thus sensitive to the facts of spatial inclusion, extension and betweenness, we would *not* wish to rule out *such* a measure function on one of our *dense* manifolds as trivial just because it assigns the *same* measure number to a closed interval of the dense manifold as to the corresponding open interval and as to the corresponding two half-open intervals. This assignment of the same number as is assigned to the closed interval is made by length functions to any of these three specified proper subintervals, which we shall call " I -triplets." Indeed, once we also impose the *second* of our nontriviality conditions as we shall do further below, *consistency* will require that each member of an I -triplet of a dense space be

¹⁴ This formulation of a weak kind of countable additivity is given in Massey's 1963 Princeton University doctoral dissertation [49]. It is to supplant the definition of weak countable additivity in his Panel article ([50], p. 333). For the latter is stated by reference to a sequence of *non-overlapping* sub-intervals. And in the case of a *discrete* space, its application to an interval $[a, d]$ containing the four points a, b, c, d in that order would call for the addition of the lengths l of the *disjoint* subintervals $[a, b]$ and $[c, d]$ to obtain the length of the interval $[a, d]$. But this addition would yield a number that does not accord with the distance assigned to the ordered pair (a, d) by a permissible non-trivial distance function, which is such that $d(a, b) = d(b, c) = d(c, d) = l[a, b] = l[c, d]$.

assigned the same measure as the corresponding closed interval. For, once the first two of our nontriviality conditions have both been imposed, the following result becomes provable, as Massey has shown: In a dense manifold, any non-trivial finitely or countably additive measure must assign the measure *zero* to every singleton of the dense manifold. And the finite additivity of the measure then assures that all three members of a given *I*-triplet are assigned the same measure as the corresponding closed interval, an assignment made by length functions. Thus, in linear continuous space one usually requires that the measures of any and all intervals be given by their lengths.

We have seen that a measure function or a length function satisfying our first necessary condition for nontriviality is sensitive to such betweenness relations as exist among the end-points of the *closed* intervals to which it pertains. And, anticipating results obtainable from the first two of our nontriviality conditions, we noted that once both of these conditions have been imposed, we are no longer free to assign different measures to the members of an *I*-triplet.

That it would be wrong to indict a measure or length function on a dense manifold as trivial merely because it assigns the same measure to the members of an *I*-triplet of that manifold as to the corresponding closed interval can be seen in yet another way. In a dense manifold, a closed interval does extend over its corresponding open interval and also over each of its two corresponding half-open intervals, the difference sets having two members and one member respectively. And although the equal measure numbers assigned to the closed interval and to the members of an *I*-triplet by a non-trivial measure do not directly render these relations of spatial inclusion or those of betweenness, these relations are nevertheless rendered indirectly by a measure satisfying our first non-triviality condition. To see this, let us first give the following definition: An interval I_2 "metrically contains" an interval I_1 with respect to a measure M , iff for *every* closed subinterval I' of I_1 , there is a closed subinterval I'' of I_2 such that $M(I'') > M(I')$. Then we can say that for any given *I*-triplet, the closed interval *metrically* contains both the open and half-open intervals though not conversely, and either half-open interval metrically contains the open interval but not conversely.

So much for the first of our necessary conditions for non-triviality of a measure-metric. It will be recalled that I stated and justified it for *P*-manifolds and *T*-manifolds. This involvement of *P*-manifolds and *T*-manifolds should *not* be taken to mean that I would refrain from similarly imposing and justifying this condition in the case of certain *arithmetic* manifolds, for example. When I now go on to state the second necessary condition for non-triviality, I shall again justify it specifically by reference to presumed facts of spatial and temporal extension. But I shall point out that depending upon the particular purposes which an *M*-metric on *other* kinds of manifolds is designed to serve, the imposition of this *second* necessary condition may or may not be desirable in the case of these other kinds of manifolds. In fact, let me develop the motivation for the second necessary condition by considering, initially, *M*-metrics on a singly extended manifold which is neither a *P*-manifold nor a *T*-manifold: the series of natural numbers. Our concern will now be with the measures which two different *M*-metrics M_1 and M_2 assign to singletons of this

manifold. If n is *any* natural number, and $\{n\}$ its corresponding singleton, let these respective measures be the following:

- (i) $M_1(\{n\}) = c$,
where c is a non-negative *constant* whose value is $1/k$ for some one positive integer k .
- (ii) $M_2(\{n\}) = n$.

Suppose that our purposes in defining a measure metric on this arithmetic manifold happen to require that the measures assigned to singletons be sensitive to their idiosyncratic arithmetical properties, properties whose articulation falls within the purview of the usual arithmetic theory of the naturals, as we noted in section 1(a). Given such purposes, these would certainly not be served by the measure M_1 above, although they would be served by the measures M_2 . On the other hand, suppose that all we are concerned to require of the measure numbers assigned to singletons is that they be sensitive to the fact that qua singletons, all singletons are *alike*, whatever their other differences might be. In that case, the constant measure M_1 might well be satisfactory. (This is, of course, *not* to say that the values which the measure function M_2 assigns to singletons are assigned *only* to singletons by that measure function.)

Hence the arithmetic manifold of naturals is homogeneous with respect to a measure function that assigns the same measure to all singletons, although that manifold is an *I*-manifold with respect to certain familiar arithmetic properties and is also metrically inhomogeneous with respect to measures $M_2(\{n\}) = n$ that make use of these properties. This co-presence of homogeneity in one respect with inhomogeneity in another should occasion even less surprise than the following fact, noted at the end of section 1(a) above: discrete denumerable *P*-space with exactly one extreme element is a homogeneous manifold with respect to metrical extension while also being an *I*-manifold with respect to spatial betweenness.

In the work of the arithmetician, *M*-metrics on his number manifolds may thus serve various alternative objectives in regard to singletons. By contrast, what is *spatially* or *temporally* pre-eminently *relevant* to the assignment of measures to the singletons of a *P*-manifold or *T*-manifold is that they *are* all singletons whose sole members are ultimate constituents having no spatial or temporal proper parts of any kind!¹⁵ Thus singletons of space and time are alike *indivisible* not merely in the sense of their set-theoretical definition that they have only one member and

¹⁵ Instants (or moments) of time are specified by simultaneity classes of events, but no class of events properly included in a particular simultaneity class is itself a simultaneity class. Thus, instants do not have proper parts that are themselves instants. Moreover, the only thing that is *temporally* relevant about the members of a given simultaneity class of events is that they stand in the relation of simultaneity to one another: In construing their simultaneity class as specifying a *single element* of the *T*-manifold, one abstracts from all of their individual differences and regards them as *temporally* indistinguishable. Thus, although simultaneity classes of events have proper subsets of events, these proper subclasses are not *temporally* distinguished from the simultaneity classes that include them, and hence they do *not* qualify as *temporal* proper parts of instants: The status of points of *P*-space and of instants of *T*-space will be discussed in greater detail in the context of relativistic space-time in section 2(c), (iv) 7.

hence do not *extend over* any non-empty set; instead, they are indivisible in the further sense that the individuals belonging to them as their sole members do *not* have proper parts in any spatial or temporal respect. Contrast this with the character of a singleton of the manifold of natural numbers such as $\{6\}$: the individual 6 which is its sole member does indeed have "arithmetic proper parts" in the sense of being the arithmetic sum of non-zero numbers. And, as we noted above à propos of the measure $M_2\{n\} = n$, an M -metric introduced by the arithmetician may indeed be designed to be sensitive to this fact about certain singletons of the manifold of naturals.

It is this further, physical sense of indivisibility of points and instants, I take it, which Russell had in mind in the passage that I quoted from him at the start of section 1(a), where he wrote ([63], pp. 66–67): "With space, . . . , if we seek for elements, we can find none except points, and no analysis of a point will find magnitudes inherent in it. . . . The magnitudes which space deals with, . . . , are relations between points. . . . There is no inherent quality in a single point, as there is in a single colour, by which it can be quantitatively distinguished from another." Since our full-fledged sense of indivisibility is a physical-*cum*-set-theoretical one, it is evident that the ascription of such indivisibility to the singletons of physical space or time is *not* equivalent to the set-theoretical tautology that all singletons have exactly one member. Massey has pointed out to me that the exceptional status of P -spatial singletons among P -spatial sets can perhaps be more vividly dramatized by means of the Leonard–Goodman language of individuals.¹⁶ Among spatial individuals certain ones are distinguished by the fact that they have no proper parts which are also spatial entities. These individuals are, of course, the *points*, and they are the counterparts of singletons among spatial sets. Other spatial individuals, such as lines and planes, are sum-individuals of two or more points and have spatial proper parts. Thus points, the individuals corresponding to singletons, are indivisible in the straightforward sense of having no spatial proper parts. And one could naturally insist that a spatial measure be sensitive to or reflect this important feature of points through the values assigned to singletons.

The sensitivity of the M -metric to our full-blown physical-*cum*-set-theoretical indivisibility of the singletons of a P -manifold or T -manifold does *not* require that the measure assigned to them must be zero. For, in a discrete P -space or T -space, that measure might, for example, be 1 or $1/k$, where k is any positive integer. Furthermore, suppose that a measure function does reflect the full-blown indivisibility of these singletons by assigning the *same* measure to all of them in *any one* kind of P -space or T -space. As is evident from the axioms for M -metrics, if that measure is zero in a continuous P - (or T -) space, this assignment fully allows that the one measure number assigned alike to all singletons of that particular space may *also* be assigned to certain of its nonsingletons. Thus, while singletons are the only sets which are indivisible in the specified sense, they need *not* be the only sets of a P -space or a T -space that are *metrically unextended* in the sense of being

¹⁶ For a discussion of the language of individuals, see [45], and [51], Section 56.3 "Postulate Systems; Calculus of Individuals."

assigned the *lowest* among the non-negative measure numbers assigned to non-empty sets of that space. In thus characterizing metrical unextendedness, I regard the attribute of metrical unextendedness as not significantly applicable to the empty set, to which every measure function assigns the measure zero.

The imposition of my first necessary condition for non-triviality was prompted and justified by the need for sensitivity to spatial and temporal facts of inclusion, over-extending and betweenness. By the same token, I maintain that a measure function on a *P*-manifold or *T*-manifold is trivial, unless it also takes cognizance of the following pre-eminent, spatially or temporally relevant fact: all singletons are *alike* indivisible in the specified full-blown sense. Assume that no specific restrictions have been imposed on the measures that may be assigned to singletons in particular, the sole restrictions being those which the axioms for *M*-metrics impose quite generally on the measures of various kinds of sets. Given this absence of any specific restrictions on the measures of the singletons of *P*-manifolds and *T*-manifolds, I maintain that the nontriviality of a measure-function requires that the sameness of singletons in regard to indivisibility be respected by suitably restricting the measures that may be assigned to them. Given our first necessary condition for nontriviality, such a restriction is imposed by a specific requirement suggested by Massey: For any one *P*-manifold or *T*-manifold, if a set *A* properly includes a singleton, then the measure $M(A)$ of *A* cannot be less than the measure assigned to *any singleton whatever*. This second necessary condition assures that the measures assigned to singletons are bounded, i.e. that they are confined to values *m* such that $0 \leq m \leq k$, where *k* is some fixed non-negative real number. For if they were not confined to a certain range of non-negative real values, *all* sets *A* that properly include a singleton and hence all nondegenerate intervals would have to be assigned the measure $M(A) = +\infty$, in contravention of our first necessary condition. For brevity, I shall speak of this second necessary condition as the requirement of "telling the spatial or temporal story" about the singletons of the respective manifolds.

Massey has pointed out that it is possible to be a good deal more specific about the range of values to which the first two non-triviality conditions confine the measures of singletons. In particular, he has shown that these two conditions suffice to require the assignment of the measure zero to any and all singletons of a *dense P*- (or *T*-) manifold; although the first two non-triviality conditions alone do not suffice to confine the measures of the singletons of our singly extended *discrete P*- (or *T*-) manifolds to a single value. Let us see how Massey reaches these conclusions. Before considering dense and discrete *P*- (or *T*-) manifolds separately, Massey points out quite generally that for finitely additive and hence also for countably additive measures, the second of our non-triviality conditions entails the following two results:

- (i) If zero is assigned to any one singleton *S*, and if any other singleton is assigned a positive finite measure *v*, then *every* singleton other than *S* must have that same non-zero measure *v*; we shall refer to this first result as "*the 0/v result*."
- (ii) If no singleton is assigned the value zero, then there must be a positive

greatest lower bound b on the measures assigned to singletons; we shall refer to this second result of Massey's as "the b -result."

Note at once that if no singleton is assigned a measure less than the positive real number b , then the sum of the measures of any *finite* number n of singletons is *not* bounded as a function of n . And in that case it follows from the axioms for a finitely additive measure that no non-degenerate closed interval of a *dense* P - (or T -) manifold can have a measure other than $+\infty$, which violates our first nontriviality condition. Hence the " b -result" just obtained tells us that our first two nontriviality conditions interdict an assignment of exclusively nonzero measures to the singletons of a *dense* manifold. In other words: Every measure function that assigns positive measure to *all* of the singletons of a dense manifold D is trivial.

This lemma must now be combined with the following further lemma of Massey's. If any singleton of a dense manifold D were to have a positive measure under a nontrivial measure function, then all the singletons of D would have to be assigned positive measures by that function.¹⁷ The combination of the latter lemma with the former then entails the following theorem of Massey's by indirect proof: A nontrivial measure, whether finitely or countably additive, must assign the measure zero to every singleton of a dense P - (or T -) manifold. By contrast, any assignment of zero to all singletons of a *discrete* manifold would violate our first nontriviality condition.

What is the upshot of our considerations so far for the measures assignable to singletons? The two nontriviality conditions we have imposed so far require that all singletons of a *dense* P - (or T -) manifold have the measure zero but allow the following latitude for the measures assigned to the singletons of a *discrete* P - (or T -) manifold: All of the singletons have positive finite measures $+m$ such that $+j \leq +m \leq +k$, where $+j$ is the positive greatest lower bound and $+k$ the positive least upper bound.¹⁸ For although the second of our nontriviality conditions allows the assignments of the measures to which the $0/v$ result pertains, these $0/v$ assignments are ruled out by the first nontriviality condition.

Before considering the imposition of a third nontriviality condition, I wish to comment on the factual content of measures satisfying our first *two* nontriviality conditions.

¹⁷ Massey has given the following proof of this further lemma of his (private communication): "Let m be a non-trivial measure, and let S be a singleton such that $m(S) = v$, where $v \neq 0$. Clearly $v \neq +\infty$, for otherwise the first non-triviality condition would be violated. So v is a positive finite real number. Suppose there is a singleton Z such that $m(Z) = 0$. Then, by the $0/v$ result, we know that the measure of every singleton other than Z is v . Let (a, c) be a non-degenerate interval of D and let b lie between a and c . Then (a, b) is also a non-degenerate interval. By the first non-triviality condition, $0 < r < +\infty$, where $r = m(a, b)$. Let k be the least positive integer such that $k \cdot v > r$, and let A be any subset of (a, b) with exactly $k + 1$ members. Then by the finite additivity of the measure, $m(A) = k \cdot v$ or $(k + 1) \cdot v$, according as Z is included or not in A . So we have $A \subseteq (a, b)$ with $m(A) > m(a, b)$, which contradicts a basic theorem about measures. Therefore, the lemma follows by indirect proof."

¹⁸ If each singleton of a discrete P - (or T -) manifold were assigned a *different* positive measure in the specified range from $+j$ to $+k$, our second nontriviality condition would restrict these assignments by the requirement that for any singletons $\{a\}$, $\{b\}$, and $\{c\}$,

$$m(\{a\}) + m(\{b\}) \geq m(\{c\}).$$

Any measure function suitably defined on any one of the kinds of *P*-manifolds and *T*-manifolds we are considering and meeting the two necessary conditions for nontriviality imposed so far will tell a spatial or temporal story which is *factual* in the following *specific respect*: With the very minor exception of a two-member interval of a *discrete P*-manifold or *T*-manifold, any set of elements of a given *P*-(or *T*-) manifold whose measure falls within the range of values assigned to singletons of that manifold shares with singletons the important property that it cannot include any non-degenerate interval. For our necessary conditions assure that, with the one noted exception, any set which includes a non-degenerate interval will be assigned a greater measure than any singleton whatever.

Some extreme conventionalists maintain that it is open to us to choose *M*-metrics for *P*-(and *T*-) manifolds which violate either or both of our nontriviality requirements, just as we are free to choose nontrivial *M*-metrics. This misleading statement is true in the uninteresting sense that we can choose to tell or not to tell the specified spatial (or temporal) story *M*-metrically. But what follows from our having the option *not* to tell the specified spatial (temporal) story? It surely does *not* follow that when we do tell it *M*-metrically, then *the specified part of the story* is true by convention! I stress that *the specific spatial (temporal) claim* above, stated by reference to nontrivial *M*-metrics, is a statement of intrinsic spatial (temporal) *fact* instead of being true by convention.

For, as will be recalled from the **Introduction** and be made much more precise and explicit in section 3 below, I do affirm the ingredience of an important conventional element in the following different claim: if two *disjoint* linear non-degenerate intervals of a *continuous P*-manifold or *T*-manifold are assigned unequal measures by a nontrivial metric m_{nt} , then the interval of greater measure contains "more space" or "more time" than the interval of lesser measure, and if two such disjoint intervals have equal measures m_{nt} , then they contain "equal portions of space or time." But what support can the *extreme* conventionalist draw from the ingredience of an important conventional element in all *M*-metrical ascriptions of spatial (temporal) inequality or equality to disjoint nondegenerate intervals in *certain kinds* of *P*-manifolds (*T*-manifolds)? Surely this ingredience does *not* detract from the factuality of the part of the spatial (temporal) story specified above, as told by *M*-metrics about these or other kinds of *P*-manifolds or *T*-manifolds. We have the choice of whether or not we tell the story of chemical composition by identifying a specimen of sulphuric acid as a specimen of H_2SO_4 or merely as, say, specimen 1. But it would be wrong to conclude from our ability to exercise this choice that if we do tell the story of chemical composition about the specimen, then what we tell is true by convention.

At the beginning of our present subsection 2 on *M*-metrics, I appealed to spatial or temporal facts of proper inclusion or extending-over among closed intervals to justify my first nontriviality condition. I now wish to give consideration to a *possible* further condition of nontriviality which is likewise *suggested* by these same facts of proper inclusion. I shall refer to this further condition as "FC". It will be formulated initially by reference to *M*-metrics, which are our current concern. But FC also has *D*-metrical, length-metrical, and *R*-metrical counterparts.

It will turn out that I shall in fact *not* impose FC on M -metrics. We shall find that the imposition of its D -metrical counterpart would involve a *non-overriding value judgment* as to the relative importance of different facets of the spatial (temporal) story. For it will emerge that there are D -metrics (on a discrete P -space with exactly one end-point) which violate the D -metrical counterpart of FC but which have the following, *perhaps* compensatory, merit: These D -metrics do tell a facet of the spatial story that might be deemed *pertinent* no less than the part told by another metric satisfying the D -metrical counterpart of FC. Although I have no comparable positive grounds for refraining from imposing FC on M -metrics, I am abstaining from doing so out of mere caution. But I am hopeful that in the case of M -metrics, others may be able to show my caution to be unfounded. Hence, I shall proceed to canvass the motivation and the consequences of imposing FC.

To lead up to the motivation for FC and to its statement, let us now consider the bearing, *if any*, of a relation of proper inclusion among two intervals on the existence of a non-zero difference between their cardinalities. Let I and J be intervals of one of our P - (or T -) manifolds M such that $(I \subset J) \cdot (I \neq J)$. If M is *discrete*, only *finite* sets qualify as intervals. In a discrete M , therefore, J 's proper inclusion of I will entail that the *cardinality* of J is *greater* than that of I . By contrast, in any one dense M , all *nondegenerate* intervals have the same cardinality, although the cardinality of a nondegenerate interval is, of course, greater than that of a singleton. Hence in any given dense M , it is *not* at all true for all intervals I, J that the proper inclusion of one of them in the other entails a difference in their cardinalities. Before Bolzano and Cantor, the concept of a hierarchy of infinite cardinal numbers was still only quite vague, if entertained at all. But historically the belief that the relation of extending-over among physical line segments *must* involve the existence of cardinally "more" points in the longer of the two segments, be the number of points finite or infinite, was a recurring theme. So was the cognate belief that the cardinality of the points of a three-dimensional region R of physical space must exceed the cardinality of points in any proper subregions of R of dimension two or one ([41], Ch. 1, and [53], pp. 83–86). And it was experienced as shockingly counter-intuitive when Bolzano ([6], pp. 41 and 42) and Cantor showed that in a dense M , the specified relations of proper inclusion do *not* involve a difference of cardinality.

Though erroneous, the unqualified intuition of which Cantor's work disabused the mathematical public contains a valid core. And this core is relevant to the requirement that measure functions tell the spatial or temporal story of proper inclusion or extending-over. The pertinent valid core is the following: There are kinds of P - (or T -) manifolds in which a difference in cardinality is *necessary* in every case of proper inclusion of one interval in another. Thus, in any discrete P - (or T -) manifold, for all intervals I and J such that $(I \subset J) \cdot (I \neq J)$, the cardinality of I is less than that of J . And in a P - (or T -) manifold in which the general property of cardinality plays this role in the proper inclusion of I in J , it may be thought that the spatial (or temporal) story of proper inclusion is told *more fully* if their measures are sensitive to their *differing* cardinalities, than if these measures satisfy merely our first nontriviality condition. For in the case of the sensitivity of the measure

function to the differing cardinalities of such intervals I and J , the measures tell us in each particular case *by how many elements* J exceeds I while extending over it rather than merely that it extends over it.

Thus, again using the requirement of telling the spatial (temporal) story, we see that at least in discrete space, it might be thought that the measures of intervals will be trivial unless they are sensitive to differences in their cardinalities. The sensitivity of a measure to *differences* in cardinality means that whenever the cardinalities are different, the measures will be different. But such sensitivity does *not* also require that whenever the cardinalities are the same, then the measures will be the same. With this *caveat*, I shall use the term “sensitivity to cardinality” as an abbreviation for “sensitivity to *differences* in cardinality.”

Let us now see whether the demand for sensitivity of the measures of intervals to their cardinalities can be formulated as a possible FC applicable not only to our discrete P - (or T -) manifolds but to the dense ones as well. And let us then consider whether such a broadly applicable condition can assume different forms in regard to the kind of function that relates the cardinalities of intervals and their measures.

That the answer to the first of these two queries is affirmative is clear from the following possible formulation: For any P - (or T -) manifold in which, for *all* I and J , the proper inclusion of one of two intervals I and J by the other entails a difference in their cardinalities, the measure $M(A)$ of any interval A must be a function of its cardinality $\bar{A}$, i.e. there must be a single valued function f such that $M(A) = f(\bar{A})$. Note that this requirement would be satisfied vacuously in our dense manifolds, since its antecedent does not apply to these manifolds. And its *hypothetical* imposition on our *discrete* manifolds in the context of the axioms for finitely additive measures would have the following logical consequence: the measure of any interval A of a discrete P - (or T -) manifold must be a *linear homogeneous function* of its cardinality $\bar{A}$, i.e. the equation $M(A) = f(\bar{A})$ takes the particular form

$$M(A) = k\bar{A},$$

where k is a non-negative real constant. In the case of singletons $\{a\}$, $\{\bar{a}\} = 1$. If k had the value zero, then the measures $M(\{a\})$ of all singletons $\{a\}$ and hence of all finite sets would be zero. And this would violate the *first* of our nontriviality conditions. Hence the constant k must be a positive real constant. And the measure of an interval would yield its cardinality upon being divided by k .

Thus, *if* the demand for the sensitivity of the measure to the cardinality were to be imposed on the intervals of one of our discrete manifolds, and if this imposition took the form $M(A) = f(\bar{A})$, then all singletons of that manifold would have to be assigned the *same* positive measure k . And, in that case, the discrete manifold in question would qualify as a “space” in the sense of our section 1(a) above. For the metrical equality of all of the singletons of a manifold was our original basis for classifying the manifold as a “space.” In the case of any one of our dense manifolds, we saw that the fulfillment of the first two of our nontriviality conditions by a finitely additive measure function suffices to qualify such a manifold as a “space”

(i.e. as metrically homogeneous) by assuring that each singleton is assigned the same measure (*zero*). And given this special usage of the term "space," it is noteworthy that in a dense denumerable space, there are no nontrivial *countably* additive *M*-metrics. For the countable additivity of the measure then tells us that all nondegenerate intervals have the measure 0, in contravention of our first necessary condition for nontriviality.

The hypothetical requirement that all singletons of a discrete manifold be assigned the same positive measure *k* would become more restrictive, if we were to demand that there be an interval of measure 1 (unit interval), i.e. that the measure 1 have a physical realization. Any interval of a discrete manifold contains an integral number *n* of elements ("atoms" or "quanta"). Hence, as I noted in an earlier publication ([28], p. 154 and 154n.), the imposition of this demand for the physical realization of the unit in discrete space requires that the measure $M(\{a\})$ of each singleton of that space satisfy the condition

$$n \cdot M(\{a\}) = 1,$$

so that

$$M(\{a\}) = \frac{1}{n},$$

where *n* is an integer whose value is determined by the choice of the unit. This means that the constant of proportionality *k* must have the value 1/*n*, so that $M(A) = 1/n \cdot \bar{A}$.

Upon combining this result $M(\{a\}) = 1/n$ with our earlier one that the singletons of a dense space must each have the measure zero, we obtain the condition on the measures of singletons which van Fraassen ([68], p. 349) attributed to me, provided that we qualify his condition (iv) there by reserving the measure zero for the singletons of dense space.

Our affirmative answer to the first of our two queries concerning a generally applicable FC now yields an important conditional conclusion in regard to the metrical homogeneity of our discrete *P*- (and *T*-) manifolds. For suppose that in such a discrete manifold the exigencies of telling the spatial (temporal) story were able to justify using the *particular* functional form $M(A) = f(\bar{A})$ to implement the putative demand for sensitivity of the measures of intervals *A* to their cardinalities $\bar{A}$. If such a justification could be given, then my argument above for assigning the same positive measure to all singletons of a discrete *P*- (or *T*-) manifold would *not* at all be an argument to the effect that we might as well assign the same measure to them, since we have no reason to do otherwise. For in the hypothetical event of such a justification, my argument for the metrical homogeneity of atomic *P*- (or *T*-) manifolds is not an argument from insufficient reason, akin to the Laplacian's invocation of the principle of indifference or equi-ignorance in defense of his *a priori* probability metric. Instead, in the presumed eventuality, my case for metrical homogeneity would rest on the existence of good and sufficient reason.

It is interesting in this connection that distance functions automatically take

metrical cognizance of the unique set-theoretical status of the ordered pair (a, a) vis-à-vis all other ordered pairs (a, b) of distinct elements a and b .

But the provision of the presumed spatial (or temporal) justification for $M(A) = f(\bar{A})$ is predicated on the feasibility of *ruling out* the alternative functional implementation $F[M(A)] = \bar{A}$ of the putative demand for sensitivity of the measure to the cardinalities of the intervals of a discrete manifold. I am greatly indebted to Gerald Massey for pointing out the existence of this alternative to me by reference to two of his examples:

(i) Let s be a variable ranging over the numbers in the geometric progression $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots$. And let each singleton of a discrete denumerable P - (or T -) manifold be assigned a *different* measure $1 + s$. Since $\sum_{i=1}^n s_i$ is less than 1 for any finite n , the finitely additive measure of any finite set A will lie between two integers l and $l+1$, where l is the cardinality of A . Hence this finitely additive measure function $M(A)$ will be sensitive to the cardinalities of intervals in the sense that the measure $M(I)$ of any interval I of the discrete manifold will uniquely specify the cardinality $\bar{I}$ of the interval, so that intervals of differing cardinality will be assigned different measures, thereby telling the spatial (temporal) story with respect to differences in cardinality. Moreover, if the measure is countably additive, any set of measure $+\infty$ will be denumerably infinite.

Clearly, in this example, the sensitivity to the cardinalities takes the weaker form $F[M(A)] = \bar{A}$ of a single-valued function from the measure to the cardinality rather than the stronger form $M(A) = f(\bar{A})$ of a function from the cardinality to the measure.

Note that on this assignment of measures $1 + s$ to singletons, no two *distinct* finite sets A and B can have the same measure, regardless of whether their cardinalities are equal or different. For if their cardinalities are the same, the fractional parts of the sums of the terms $1 + s$ will be different. And for sets of different cardinality, the integral parts of the sums will be different, while the fractional parts are each less than 1 and cannot compensate for that integral difference. Hence in the case of this "pathological" finitely additive measure function M , equality and inequality of finite measures go hand-in-hand respectively with the (logical) identity and nonidentity of sets. Thus, here M -equality generates the same partition of the class of intervals as does the equivalence relation of identity: The quite uninteresting partition in which each equivalence class of intervals contains only a single member! By contrast, in the case of our "nonpathological" measure function $M(A) = 1/n \cdot \bar{A}$, equal measures are assigned to any two nonidentical sets of the same cardinality, and equal measures betoken equi-cardinality for distinct sets as well, while unequal measures betoken not only logical nonidentity of sets but the inequality of their cardinalities.

(ii) Massey's second example of a measure function on a discrete manifold whose sensitivity to the cardinality takes the weaker form $F[M(A)] = \bar{A}$ is the following: All but one of the singletons of the discrete manifold have the measure 1,

and the remaining singleton receives the measure $1 + \frac{1}{3}$. The measure of any finite set of l elements will now be either l or $l + \frac{1}{3}$, and if the measure of a set is either l or $l + \frac{1}{3}$, its cardinality will be l . In this case, two sets will have the same cardinality if their measures either are equal or differ absolutely by $\frac{1}{3}$, and two sets will have different cardinalities if their measures differ absolutely by more than $\frac{1}{3}$.

The import of Massey's two examples of "pathological" measures on a discrete manifold for our formulation of the FC is the following: The FC can implement the putative demand that the measures of intervals be sensitive to their cardinalities by imposing the weaker requirement $F[M(A)] = \bar{A}$ rather than the stronger one $M(A) = f(\bar{A})$ ingredient in our *tentative* first formulation above. And since the justification for our actual or hypothetical non-triviality conditions derives solely from the need to tell the spatial (temporal) story, the *putative* imposition of FC would take only the following parsimonious form:

(FC) For any P - (or T -) manifold, and for all *intervals* of that manifold, there must be a function F from the measure of an interval to its cardinality.

Note that if a finitely or countably additive measure function on a continuous P - (or T -) manifold satisfies our first two nontriviality conditions, then it will likewise satisfy this FC. For in that case, any *interval* that has the measure zero will be a singleton, and any interval having a positive measure will have the cardinality of the continuum. In the case of a dense *denumerable* P - (or T -) manifold, we saw earlier that there are no countably additive measures that satisfy even the first two of our nontriviality conditions, let alone FC. For if a nontrivial measure on that kind of manifold were countably additive, then all intervals, degenerate and non-degenerate, would have the measure zero. The reason is that singletons of a dense P - (or T -) manifold must have the measure zero even if a merely finitely additive measure function on such a manifold is to be nontrivial.

But, as Massey has shown, there are finitely additive measure functions on a dense denumerable P - (or T -) manifold which satisfy our two nontriviality conditions as well as FC.¹⁹

After we have stated our reason for *not* imposing FC later on, we shall return to Massey's first example of a pathological measure in order to consider the imposition of *another* candidate for being a third nontriviality requirement. But first we must deal with such counterparts of our first two nontriviality conditions as need to be imposed on metrics other than M -metrics.

3. Length functions (" L -metrics"). I motivated and justified our nontriviality conditions for M -metrics by the requirement of telling the spatial (or temporal) story. Hence I can now invoke the same justification to state very briefly what counterparts of these nontriviality conditions are to be imposed on length functions

¹⁹ Private communication. Massey has shown how to construct a finitely additive measure on the ring R of subsets of the set S of *rational* points of the Euclidean line, where P is the set of bounded "intervals" of S , and R is the set of all finite unions of mutually disjoint members of P . Massey's measure is constructed in a way similar to that in which Halmos's measure $\bar{\alpha}$ is defined ([34], pp. 35–36).

("L-metrics"), i.e. on the length $L(x)$ of closed intervals x . The concept of length involved here is *not*, of course, the generalized one of *indefinite* R -metrics.

The first nontriviality condition is taken over *mutatis mutandis*. And then the fulfillment of the requirement $L(x) \neq L(y)$ in that condition will assure that $L(y) > L(x)$, since the length function was explicitly assumed to be countably additive in Massey's weak sense on physical grounds (see p. 477).

Let us recall the discussion of the connection between length and distance in section 2(a), subsection 3 above. Clearly, the length $L(\{a\})$ of any singleton $\{a\}$ is zero, as is the distance $d(a, a)$, and the length of any nondegenerate interval $[a, b]$ must be positive, being the sum (or integral) of positive (infinitesimal) lengths. But at least in a continuous multiply extended space, the corresponding positive sum (integral) of positive (infinitesimal) distances cannot be equal to the distance (a, b) between a and b , unless the interval $[a, b]$ is a geodesic path with respect to the length metric ds or is itself infinitesimal. Note that Massey's *weak* kind of countable additivity (cf. section 2(b), subsection 2 above) does *not* render the *lengths* of *singletons* additive and hence does not issue in inconsistency with the positive lengths of nondegenerate intervals in either discrete or dense denumerable space. Anyhow, the properties of length functions render it superfluous to impose on these functions a counterpart to our second nontriviality condition governing M -metrics.

If FC were to be imposed on M -metrics, then one would also impose the counterpart of FC and require the specified sensitivity of the length function to the cardinality.

Having drawn these parallels between the nontriviality conditions governing M -metrics and L -metrics, it will be useful to comment briefly on the relation between these metrics with respect to the intervals of some of our P - (or T -) manifolds. If such a manifold is continuous, the conditions we have laid down allow us to follow the customary practice of requiring that the measures of any and all intervals be given by their lengths. And in the case of a dense denumerable P - (or T -) space, for any nontrivial length function, there is at least one nontrivial finitely additive measure function that assigns the same numbers to any interval as that length function. Hence in the case of dense denumerable space, we can allow that the nontrivial finitely additive measures of any and all intervals be given by their lengths, provided that we require lengths to be only *weakly* countably additive in Massey's sense. For if we were to require (on physical grounds) that lengths be countably additive in the usual stronger sense, then there would be no such length functions in dense denumerable space, and also no finitely additive length functions in discrete space.

In the case of a discrete P - (or T -) manifold, we saw that no nontrivial measure can assign zero to all singletons. (Indeed, we noted that even the assignment of zero to only one singleton as part of the assignments of Massey's $0/v$ result is ruled out by our nontriviality conditions.) By contrast, the *length* of any singleton is *zero*. Therefore, in our *discrete* P - (or T -) manifolds, the nontrivial measures of intervals cannot be equal to their lengths. Instead, here length and measure of intervals are related in other ways, if at all. Let us confine ourselves to merely

illustrating these other ways by reference to (1) our measure function $M(A) = 1/n \cdot \bar{A}$ for the case of sets A which are intervals I , and (2) the one "pathological" measure concocted by Massey which is given by

$$M(A) = \bar{A} + \epsilon,$$

where ϵ is the sum of all the values s corresponding to the particular singletons included in A , and is between 0 and 1.

(1) If $M(A) = 1/n \cdot \bar{A}$, a permissible relation between these measures and non-trivial lengths of intervals I of discrete space is given by the linear equation

$$L(I) = M(I) - \frac{1}{n}$$

or

$$L(I) = \frac{1}{n}(\bar{I} - 1).$$

Note that if our choice of unit measure corresponds to $n = 1$, i.e. if a singleton is assigned the measure 1, the length of any interval, and thus also the distance between its end-points, is obtained by subtracting 1 from its cardinality. And a singleton is assigned the length zero, as required, no matter what our choice of a unit of measure.

(2) Now consider Massey's "pathological" measure

$$M(A) = \bar{A} + \epsilon,$$

where the variable ϵ is confined to values $0 < \epsilon < 1$. Being mindful of the fact that the length of a singleton must be zero, we note that a permissible relation between these measures and nontrivial lengths of intervals is given by the following linear function of two variables $M(I)$ and ϵ :

$$L(I) = M(I) - (1 + \epsilon).$$

But, of course, this equation reduces to

$$L(I) = \bar{I} - 1,$$

which is the same as the one obtained in case (1) for $n = 1$. Hence even the use of a "pathological" measure function is compatible with the use of a "normal" length function.

4. *D*-metrics. We have noted the intimate relation between the familiar physical determination of distance and of length, as well as some of the complexities of that relation in certain multiply extended manifolds. More generally, in any kind of *P*-space or *T*-space, the physical interest of assigning distance numbers to pairs of points derives from the fact that they are the termini of (one or more) *intervals* of specified lengths or measures! The intimacy of the relation between length and distance makes for the fact that if a distance function is to tell the spatial (temporal) story in a context in which length is also defined, the nontriviality conditions governing the corresponding lengths of certain intervals become pertinent. Our present

concern is to state nontriviality conditions for distance. But mindful of the aforementioned complexities of the connection between length and distance in multiply extended manifolds, we shall content ourselves here with referring to that connection to state nontriviality conditions for distance functions on a *singly* extended manifold and the D -metrical counterpart of FC for such a manifold. It is to be clearly understood, however, that the connection between distance and length would likewise make itself felt in any statement of nontriviality conditions governing a distance function on a multiply extended manifold for which length is also defined.

The D -metrical counterpart of our first nontriviality condition for length demands sensitivity to spatial (or temporal) betweenness as follows: For any three distinct points a , b and c , if b is between a and c , then $d(a, b) < d(a, c)$. It will be recalled that measure functions and length functions satisfying their first nontriviality condition are likewise sensitive to betweenness.

In our singly extended P - (or T -) manifolds, the distance between any two points a and b is given by the length of the closed interval $[a, b]$ determined by a and b . Indeed, here no less than in a multiply extended P -manifold, we again invoke the fact that the physical interest of assigning distance numbers to pairs of points derives from their being the termini of (one or more) *intervals* of specified lengths or measures. Thus the equality of length and distance in a singly extended P - (or T -) manifold on which a length function is also defined should be regarded as a condition necessary for the spatial (temporal) relevance of the mathematical distance function.

If one were to demand that the length function be sensitive to the cardinality by analogy to FC, then consistency would require us to demand that there be a function from the distance between any two points a and b to the cardinality of the closed interval $[a, b]$. Since the latter interval has exactly two end-points, there would also be a function from the distance $d(a, b)$ to the cardinality of those points that are between a and b (in the sense of Huntington). Just as in the case of measure functions and length functions, the imposition of this requirement on distance functions might be thought to be especially germane to telling the spatial (temporal) story with respect to a *discrete* P - (or T -) manifold. But the sensitivity of the distance function to the cardinality might be deemed necessary for that metric's telling of the spatial story of discrete space, even if no length function were defined on its intervals but only *measures* (an M -metric).

Let us recall the stated equality of the length of a closed interval and the distance between its end-points in singly extended P - (or T -) manifolds on which a length function is defined. Then our statements concerning the mathematical relation between length and measure are seen to apply, *mutatis mutandis*, to the relation between distance and measure. For the case of a discrete space, this means that if $M(I) = \bar{I}$, then

$$d(a, b) = M[a, b] - 1$$

is an allowed relation, as pointed out in van Fraassen's Panel paper ([68], p. 350) in explication of what I had written earlier ([28], p. 147).

5. Indefinite R -metrics of relativity theory. In our discussion of the *null* geodesics of Minkowski space-time in section 2(a), we saw that all intervals on the world-lines of light rays have the same vanishing space-time "length." By contrast, it is clear from our definition there of the metric $d\tau$ that on *time*-like geodesics of that space-time (e.g. on the world-line of a material clock or particle), the non-negative real-valued space-time "lengths" $d\tau$ (or τ) will satisfy the following condition: For all intervals I and J on time-like geodesics such that $(I \subset J)$, if the difference set $J - I$ includes a nondegenerate interval, then the space-time "length" of J is greater than that of I . Clearly, the zero space-time lengths of closed intervals on *null* geodesics do *not* satisfy the counterpart of our first necessary condition for the nontriviality of M -metrics. But the latter counterpart condition is satisfied by the non-negative real-valued space-time lengths of intervals on *time*-like geodesics.

Very important considerations, including highly *nontrivial* claims of physical invariance, prompted the theory of relativity to describe the space-time manifold as a Riemannian space with an *indefinite* metric having the features just stated. In *this* context, it would be wrong-headed to indict the indefinite relativistic metrics as trivial just because *there are* world-lines on which they fail to satisfy the counterpart of our first necessary condition for the nontriviality of M -metrics. For this failure to tell R -metrically the story of the relations of proper inclusion among intervals on null geodesics is the price that is being paid for providing other compensatory information of great physical interest in a mathematically manageable way!

Although the above indefinite R -metrics violate the counterpart of the strong kind of first nontriviality condition for M -metrics, they do satisfy a weaker one. To state the latter, consider the universe of discourse of all linear closed intervals of the space-time (PT -manifold) on which an indefinite relativistic R -metric $d\tau$ is defined. Denote the property of being a *nonsingleton* by " NS ." Then the weaker kind of nontriviality condition which is satisfied by this type of R -metric can be stated as follows:

$$\sim(I)(J)\{[NS(I) \cdot NS(J) \cdot (I \subset J) \cdot (I \neq J)] \supset [d\tau(I) = d\tau(J)]\}.$$

Furthermore, by assigning zero to *all* singletons, the indefinite R -metrics assign the *same* space-time "length" to all singletons, thereby satisfying the counterpart of the requirement which we imposed on the M -metrics in the case of a dense P - (or T -) manifold. But by contrast to the behavior of nontrivial M -metrics, there is an infinite class of nondegenerate intervals to each of which the indefinite R -metrics assign the same "length" as they assign to singletons. For precisely this reason, these metrics violate the counterpart of the hypothetical FC.

6. Conclusions on nontriviality. We have considered such counterparts of our first two nontriviality conditions and of FC as would be appropriate to metrics other than M -metrics. And we are now ready to consider the reasons which prompted me to refrain from imposing FC as a third nontriviality condition. We stated the proposed FC in the context of M -metrics and then referred to the L -metrical, D -metrical and R -metrical counterparts of FC. I shall now give reasons for *not* imposing the D -metrical counterpart of FC. And in view of the stated inter-

connections between D -metrics and L -metrics, these reasons have relevance to L -metrics as well. Furthermore, it seems to me the better part of caution to hesitate to impose FC even on M -metrics, although I have no positive grounds for this hesitation comparable to those which I am about to state by reference to D -metrics.

In section 1(a), we saw that in a discrete P -space with one extreme element (i.e. in a discrete P -space which is also an " E -space"), the betweenness relation on that space permits us to individuate all of its points by characterizing the points as the first, second, . . . k^{th} , $(k + 1)^{\text{st}}$, etc. Thus, if we take an ordered pair (a, b) of points the first of which is the a^{th} point and the second of which is the b^{th} point, we can define a D -metric on that E -space as follows:

$$D(a, b) = |b^2 - a^2|,$$

where $|b^2 - a^2|$ is the absolute value of the specified difference. Since a and b are always both positive, we can also write

$$D(a, b) = (b + a) |b - a|.$$

And if $\bar{I}$ is the cardinality of the interval $[a, b]$, this becomes

$$D(a, b) = (b + a)(\bar{I} - 1),$$

which is a function of both the cardinality and $(b + a)$. We see that whenever both members of an ordered pair (c, d) of points precede both members of an ordered pair (e, f) while the cardinalities of the intervals $[c, d]$ and $[e, f]$ are the same, it will be the case that

$$D(c, d) < D(e, f).$$

Therefore, for point pairs corresponding to intervals of equal cardinality, this metric is sensitive to the *ordinal proximity* of a point pair to the one extreme point of the given discrete P -space. And such sensitivity might be deemed of spatial interest on a par with the spatial interest of the sensitivity to cardinality demanded by the D -metrical counterpart of FC. But since our metric $D(a, b)$ here makes $\bar{I}$ a function of *both* $(b + a)$ and $D(a, b)$, this D -metric does *not* satisfy the counterpart of FC. I know of no *overriding* reason to consider sensitivity to cardinality more informative spatially in this context than the stated sensitivity to ordinal proximity to the end-point. Hence I hesitate to impose the D -metrical counterpart of FC as a condition of nontriviality. And, as mentioned above, caution leads me to hesitate as well in regard to the imposition of FC itself on M -metrics.

That it is only caution in the case of M -metrics can be seen by reference to those species of *discrete* P -spaces (or T -spaces) which have one extreme element (" E -space") or two extreme elements (" F -space"). Consider such an E -space first. And recall Massey's first example of a "pathological" M -metric, which calls for the assignment of a *different* measure $1 + s$ to each singleton, where s is a variable ranging over the numbers in the geometric progression $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots$. Clearly, in our E -space these measures can be assigned such that the k^{th} singleton of the E -space receives the measure $1 + 1/2^k$. In that case, Massey's "progression" measure

function will be *ordinally* sensitive in addition to satisfying FC! And I can think of no ordinal part of the spatial story of our discrete *E*-space that is left out by this metric and whose rendition might require an *M*-metric that *violates* FC.

In the case of a discrete *P*-space which is an *I*-space containing at least five points, the singletons respectively containing the two extreme points can each be assigned a special measure (say 1.5) for *ordinal* reasons while all other singletons can be assigned the measure 1. This would have the result that at least one interval would be assigned the measure 3, as would the set containing just the two end-points. Thus, there would be no function from the measures of *sets* to the cardinality. But since the set containing just the two end-points is not an interval, this state of affairs would not violate FC, which pertains only to intervals.

The reasons for having adopted the first two of our nontriviality conditions seem to me more clear-cut as compared to those warranting the adoption of FC and of its *D*-metrical counterpart. Hence I see no basis for feeling ambivalent about the fact that our first two conditions of nontriviality have the following consequence: In any *dense* space having one or two extreme points (i.e. in any dense "*E*-space" or "*F*-space"), the special ordinal role of the extreme points *cannot* be marked by assigning them measure numbers differing from those assigned to all other singletons. For we saw that our first two conditions entail that the measure of any singleton of a dense space must be zero.

It remains to consider one other candidate for being a third nontriviality condition. We saw that in a discrete *E*-space, Massey's pathological progression *M*-metric can be rendered ordinally sensitive in addition to automatically being cardinally sensitive in *any* discrete denumerable space. But we also noted earlier that no two *distinct* finite sets can receive the same measure from this *M*-metric: *M*-equality and inequality of finite measures go hand-in-hand respectively with the (logical) identity and nonidentity of sets. In particular, here *M*-equality generates the same uninteresting partition of the class of intervals as does the equivalence relation of identity.

It might be urged that identity is a trivial kind of spatial equality relation and that to have a nontrivial spatial equality relation, at least two nonidentical intervals should be *M*-equal. It will be recalled that a measure assigning the same value 0 to all intervals is trivial, because it issues in universal *M*-equality among distinct intervals, thereby violating our first nontriviality condition. By the same token, it might be held that a measure which assigns different values to every two distinct intervals is likewise trivial, because it issues in universal *M*-inequality among distinct intervals. From this point of view, if *x* and *y* are intervals, it might be proposed that we lay down the following condition as necessary for nontriviality

$$(\exists x)(\exists y) [(x \neq y) \cdot M(x) = M(y)].$$

Analogously in the case of *D*-metrics, let *x* and *y* each be ordered pairs whose members are *distinct* from one another. Then it might be proposed that any nontrivial *D*-metric must satisfy the analogous condition.

But it seems to me that this proposed third necessary condition will not pass muster by the criterion of triviality that we have invoked all along: failure to tell

specifiable parts of the spatial or temporal story. For we saw that Massey's progressive M -metric is not only cardinally informative but has the capability of being ordinally informative as well when relevant, i.e. in discrete E -space. And I cannot think of any part of the spatial story that is left out by this M -metric, even though it has the feature of assigning different measures to any two distinct intervals and even to any two distinct *sets*.

Hence if the latter feature were to constitute grounds for indicting this M -metric as trivial, it would seem to do so on the strength of a sense of triviality wider than the one I am prepared to countenance. Yet here no less than in the case of FC, others may be able to show that I am overlooking something which does warrant the imposition of the requirement that at least two distinct intervals must be M -equal.

Hence Massey's example of the progressive "pathological" measure function on a discrete P - (or T -) manifold seems to me to be nontrivial in its own right and also to differ nontrivially from those nontrivial M -metrics on such a manifold which are of the form $M(A) = 1/n \cdot \bar{A}$. (Of course, the latter nontrivial M -metrics differ from one another only trivially, i.e. by a multiplicative constant.) Thus, there are significantly different kinds of nontrivial M -metrics on a discrete P - (or T -) manifold: Any M -metric of either kind tells the story of spatial (temporal) proper inclusion of intervals and indivisibility of singletons, as well as of cardinality, and yet the two kinds of M -metrics differ in their ascription of M -metrical equality. In section 3 below, I shall consider the bearing of this kind of alternative M -metrizability of a discrete P - (or T -) manifold on the following question: Are the M -metrics of a discrete P - (or T -) manifold on a par in regard to conventionality with the nontrivial M -metrics on a continuous or dense denumerable P - (or T -) manifold? And my answer will be emphatically in the *negative*.

One of the important reasons for our discussion here of nontrivial metrics is that our subsequent concern with the existence of intrinsic metrics on certain kinds of manifolds will focus on *nontrivial* intrinsic metrics.

(c) *Intrinsicness of a metric to the intervals of a manifold, to ordered pairs of elements of a manifold, and intrinsicness to the manifold as a whole*

(i) *Preliminary remarks on the underlying conception; its bearing on a space-theory of matter*

In earlier writings ([24], pp. 8–10, and [28], pp. 147–149, and Ch. III, sections 2.9–2.12 inclusive), I referred to Riemann's Inaugural Dissertation and to H. Weyl's gloss on it ([71], pp. 97–98) as the basis for my prior adumbrations of the distinction between intrinsic and extrinsic metrics. And I claimed that while Riemann and Weyl held discrete P -space or granular P -space (be it singly or multiply extended) to be endowed with a *nontrivial* intrinsic metric, they asserted continuous P -space to be devoid of any such metric. Furthermore, I attributed to Riemann and Weyl the conception that *metrically*, continuous T -space and PT -space are intrinsically amorphous in the same sense as continuous P -space.

The aim of the present Section is to replace my prior adumbrations of the intrinsicness and extrinsicness of metrics by considerably more precise formulations. To set the stage for doing so, I shall give a fuller exegesis of Riemann and

Weyl than heretofore. This exegesis will exhibit the inspiration for the intuitions underlying the more systematic formulations which are to follow. And it will likewise serve to document my interpretation of Riemann and Weyl in reply to critics who have impugned it.

In his Inaugural Dissertation, Riemann devoted Section I to "The Concept of an n -ply Extended Manifold." And in that Section I, subsection I, he wrote:

Determinate parts of a manifold, distinguished by a mark or by a boundary, are called quanta. Their comparison as to quantity comes in discrete magnitudes by counting, in continuous magnitude by measurement. Measuring consists in superposition of the magnitudes to be compared; for measurement there is requisite some means of carrying forward one magnitude as a measure for the other. In default of this, one can compare two magnitudes only when the one is a part of the other, and even then one can only decide upon the question of more and less, not upon the question of how many [much] ([61], p. 413).

Riemann's concluding Section III, which is entitled "Application to Space" begins with the following words:

Following these investigations concerning the mode of fixing metric relations in an n -fold extended magnitude, the conditions can now be stated which are sufficient and necessary for determining metric relations in space ([61], p. 422).

And in the last subsection of that concluding Section III, Riemann wrote:

... in the question concerning the ultimate basis of relations of size in space ... the above remark is applicable namely that while in a discrete manifold the principle of metric relations is implicit in the notion of this manifold, it must come from somewhere else in the case of a continuous manifold. Either then the actual things forming the groundwork of a space must constitute a discrete manifold, or else the basis of metric relations must be sought for outside that actuality, in colligating forces that operate upon it ... ([61], pp. 424-425).

Weyl introduces his own citation of the full passage in Riemann's concluding remarks from which my last quotation is taken ([71], p. 97) by saying that the "full purport" of "the concluding remarks of his essay ... was not grasped by his contemporaries. ..." And then Weyl says of these concluding remarks of Riemann's:

To make them quite clear I must begin by remarking that Riemann contrasts *discrete* manifolds, i.e. those composed of single isolated elements, with *continuous* manifolds. The measure of every part of such a discrete manifold is determined by the *number* of elements belonging to it. Hence, as Riemann expresses it, a discrete manifold has the principle of its metrical relations in itself, *a priori*, as a consequence of the concept of number ([71], p. 97).

And immediately following his full citation of the last three paragraphs of Riemann's Section III, 3, Weyl writes:

If we discard the first possibility, "that the reality which underlies space forms a discrete manifold"—although we do not by this in any way mean to deny finally, particularly nowadays in view of the results of the quantum-theory, that the ultimate solution of the problem of space may after all be found in just this possibility—we see that Riemann rejects the opinion that had prevailed up to his own time, namely, that the metrical structure of space is fixed and inherently independent of the physical phenomena for which it serves as a background, and that the real content takes possession of it as of residential flats ([71], pp. 97-98).

Our explication below of "intrinsic," as applied to certain kinds of metrics, will

be addressed to Riemann's use of the term "implicit" ("*schon enthalten*") ([60], p. 286): For the case of a *discrete* manifold, Riemann characterized "the ultimate basis of relations of size in space" and "the principle of metric relations" as "implicit in the notion of this manifold." And Weyl employed the phrase "in itself" to convey the same idea when he said on Riemann's behalf that "a discrete manifold has the principle of its metrical relations in itself, *a priori*, as a consequence of the concept of number." Likewise, our explication below of "extrinsic," as applied to certain kinds of metrics, is inspired by Riemann's conception of a manifold for which "the principle of [non-trivial] metric relations . . . must come from somewhere else" and "must be sought for outside that actuality," because that principle is *not* "implicit in the notion of this manifold."

For the latter, Riemann's case in point is spatial measurement in continuous P -space for which, he tells us, "there is requisite some means of carrying forward one magnitude as a measure for the other."

As we saw in our citations above, immediately before saying this, Riemann characterizes measurement in a *continuous* manifold in his Section I on "The Concept of n -fold Extended manifold," by saying that "Measuring consists in superposition of the magnitudes to be compared." Hence we must be mindful hereafter of the following very sound observation of Russell's ([63], p. 66): "The fact is, Riemann had [P -] space in his mind from the start, and many of the properties which he enunciates as belonging to all manifolds, belong, as a matter of fact, only to space."

Why does Riemann reason that in the case of continuous P -space, the basis of nontrivial metric relations "must come from somewhere else" or from "outside" that manifold altogether? It would seem clear that he draws this conclusion from the premiss that there does not exist *any* such basis which is "inside" that manifold M or "implicit" in it. The *weaker* premiss that the *topology implicit in M* does not furnish a basis for metric relations would surely not have sufficed to sustain the inference that such a basis "must come from somewhere else" in the case of M , unless it were distinctly claimed as well that only topological properties are implicit in M . And if Riemann did make this further assumption tacitly in the case of continuous P -space, he certainly did *not* do so on the strength of having simply *defined* the "implicit" properties as the topological ones. Indeed, if we are mindful of Russell's important comment on Riemann's meaning, then we see that Riemann's actual thought—as distinct from some of his unguarded formulations—gives no grounds for supposing that he intends the phrase "implicit in the notion of continuous P -space" to be synonymous with "implicit in the notion of the *continuity* of P -space." More generally, I interpret Riemann as having intended the phrase "implicit in the notion of [P -] space" to be intensionally *nonsynonymous* with "implicit in the notion of the *topology* of P -space." Although the continuity of a manifold is itself a topological property of that manifold, the properties "implicit" in a continuous P -space are *not* confined by *definition* to the *topological* ones. Thus, I take Riemann to have *allowed* for the existence of a nonempty set of properties which are implicit in continuous P -space, although these implicit properties

are not topological properties of that space. Hence, I believe that his construal of "implicit" was such that the assertion

- (I) Be it topological or nontopological, there exists no kind of *implicit* (intrinsic) basis for nontrivial metric relations in continuous P -space

is *stronger* than the assertion

- (II) The *topology implicit* in continuous P -space does not constitute a basis for nontrivial metric relations on that manifold.

And I believe furthermore that Riemann postulated the stronger of these two assertions, i.e. that he assumed (I) as a premiss, when he inferred for continuous P -space that the basis of metric relations "must be sought for outside that actuality" altogether. Hereafter, I shall refer to this stronger assumption (I) as "Riemann's Metrical Hypothesis" or as "RMH." In some earlier writings, I was concerned to expound the philosophical import of RMH in regard to the ontological and epistemic status of the ascription of a metric geometry to continuous P -space. And after articulating the meaning of "implicit" (i.e. "intrinsic") appropriate to RMH, stating it more precisely and commenting on its credentials in the present section 2(c), I shall endeavor anew to develop its philosophical import in section 3 below.

An alternative way of stating RMH is to say that continuous P -space is intrinsically metrically amorphous. On earlier occasions, I have claimed that the assumed truth of RMH necessarily makes for an important conventional ingredient in the ascriptions of metrical equality and inequality to the spatial intervals as generated by *any* nontrivial metric on continuous P -space. And I have maintained that no corresponding conventional ingredient inheres in the *cardinality-based* metrical equalities of the intervals of a discrete P -space. For reasons that will become clear below, I did indeed adduce topological considerations pertaining to continuous P -space as *inductive* support for RMH, which is an *empirical* hypothesis. And unfortunately, I carelessly failed to state that I viewed this support as only inductive. Thus, I did *not* wish to be understood as claiming that the mere non-existence of an intrinsic *topological* basis for metric relations in continuous P -space entails the presence of the specified conventional ingredient. For, as I just explained, I read Riemann to the effect that the mere nonexistence of an intrinsic topological basis does not necessarily preclude the existence of a *nontopological* one.

Indeed, I was aware of the fact that there are (linear) continuous manifolds whose elements are of inhomogeneous constitution, but which possess *intrinsic* metrics ([28], p. 261; [24], pp. 16–17) even though the intrinsic properties of the intervals, on which these intrinsic metrics are based, *fail* to be preserved under at least one homeomorphism. Thus, consider the arithmetic continuum of real numbers r . The transformation

$$r \leftrightarrow r^3$$

qualifies as a topological one, being biunique and bicontinuous. Note that if a and b are any two real numbers, then the absolute value $|b - a|$ of their *difference* is preserved under the homeomorphism $r \leftrightarrow r^3$. Nor does this homeomorphism

preserve the *equality* among the absolute values of the differences $|b - a|$ corresponding to two different intervals in the real number continuum. Yet, that non-negative difference $|b - a|$ can generate the following metrics, each of which will turn out to qualify as intrinsic by our definitions below:

(i) the countably additive M -metric $M[a, b] = |b - a|$,

and

(ii) the D -metric $D(a, b) = |b - a|$.²⁰

Being aware of the existence of intrinsic yet *nontopological* properties in the case of the real number continuum, I was only inductively inferring RMH as an empirical hypothesis (regarding both P -space and T -space) but *not deducing* it, when I wrote: "Accordingly, the continuity we postulate for physical space and time furnishes a *sufficient* condition for their intrinsic metrical amorphousness" ([23], p. 413 and [28], p. 13). In particular, I was *not* deducing RMH from topological considerations, coupled with the claim that the elements of continuous P -space (and T -space) are of homogeneous constitution. And by the same token, I was *not deducing* any claim of conventionality from these premisses.

Thus, I did *not* argue, as Demopoulos [13] represents me to have done, that "the congruence of spatial intervals is nonconventional only if (a') . . . it is capable of being based on some nonconventional, monadic property of the points contained in the intervals; or (b') it is capable of being based on cardinality or some other topological property of the intervals." In earlier parts of section 2 above, we noted the intimate though sometimes complicated physical relation between length, which is defined on spatial *intervals*, and distance, which is defined on ordered *pairs* of points. Thus, the inductive support which topological considerations provide for RMH pertains to a construal of RMH as applying to D -metrics no less than to L -metrics. And, by the same token, the bearing which RMH has on the presence of a conventional ingredient in metrical ascriptions of equality and inequality does not depend on whether these metrical ascriptions are generated by D -metrics or by L -metrics. It is therefore both odd and unavailing for Demopoulos to criticize the argument which he incorrectly imputed to me by saying that "it omits the most important fact about a spatial interval: that it is connected with the *relation* of distance."

Demopoulos believes that there are distance functions on continuous P -space which effect a wholly *nonconventional* (he calls it "real") and unique partition of its intervals into a set of equivalence classes of congruent intervals. The status he ascribes to these equivalence classes is held *not* to be predicated on any choice of a conventional transported standard of length-equality. I have, of course, never denied that *if* continuous P -space does provide an intrinsic basis for a certain class

²⁰ Note that the homeomorphism $r \leftrightarrow -kr$ (where $k > 1$) preserves neither the absolute value $|b - a|$ nor the relation $>$ between r_2 and r_1 . Thus the binary property $>$ does *not* qualify as a topological one in the technical sense of mathematical topology even though it is a *non-metrical*, ordinal property. Yet in the literature of the philosophy of science, the term "topological" is often used so as to have only the weak meaning of "non-metrical."

of D -metrics whose members differ from one another only by a multiplicative constant, then the latter putative intrinsically-based D -metrics would single out a unique, nonconventional partition of point pairs. But regrettably, Demopoulos does not supply even a hint as to *which* class of only trivially different D -metrics $ds^2 = g_{ik} dx^i dx^k$ on an n -dimensional continuous P -space has this distinguished status and why.

This brings me to the question whether Riemann's denial of the existence of nontrivial intrinsic metrics on continuous P -space may be construed, as Earman has suggested [16], to assert *no more* than that the metric of continuous P -space (and of continuous PT -space) "is a dynamical object whose behavior is not independent of the physical processes which perform in the space-time arena." On Earman's proposed reading of Riemann, there would be no warrant for attributing the claim which I have called "RMH" to Riemann. And this would then undercut my invocation of Riemann's metrical philosophy as a basis for the following thesis: There is an important conventional ingredient in the metrical equalities and inequalities generated by any nontrivial metric on the intervals of continuous P -space.

Riemann's advocacy of a dynamical conception of the metrical structure is contained in the following two consecutive sentences, only the first of which we had occasion to cite above:

Either then the actual things forming the groundwork of a space must constitute a discrete manifold, or else the basis of metric relations must be sought for outside that actuality, in colligating forces that operate upon it.

A decision upon these questions can be found only by starting from the structure of phenomena that has been approved in experience hitherto, for which Newton laid the foundation, and by modifying this structure gradually under the compulsion of facts which it cannot explain ([61], p. 425).

Taken in the context of the remainder of our several quotations from Riemann, it seems to me that this passage sustains the following set of interpretations as against the narrow one put forward by Earman:

1. As noted plainly by Weyl, one of the things that Riemann is clearly telling us about P -space is this: If "the actual things forming the groundwork of a space . . . constitute a discrete manifold" rather than a continuous one, then there is no corresponding problem at all of going "outside that actuality" to seek the dynamical ground for the CARDINALITY-BASED metric relations among its intervals, metric relations which are "implicit". This is *not* to say that Riemann is thereby precluding the possibility of explaining why the spatial manifold is discrete to begin with. Nor am I construing him to be gainsaying the feasibility of *extrinsically* metrizing a discrete P -space by some *nonimplicit* and nontrivial metrics or other that generate metrical equalities and inequalities which are *incompatible* with the cardinality-based ones! Nor yet am I interpreting Riemann as saying that in a discrete P -space the nonimplicit metrical relations corresponding to some such extrinsic metrization could not conceivably have a dynamical basis. For none of these possibilities impugn the tenability of the following claim of which RMH is merely a part: Unlike continuous P -space, discrete P -space is endowed with (at least) one set of

nontrivial metrical relations that are implicit and guaranteed to be independent of contingent dynamical factors (initial or boundary conditions).

2. The P -space of Newtonian physics is a continuous manifold whose elements are of homogeneous constitution. Contrary to the Newtonian metrical philosophy of his contemporaries, Riemann held that *unlike discrete P -space*, continuous P -space is devoid of fixed "residential flats," as Weyl puts it. As shown by Riemann's text, it was from precisely *this* premiss that he drew the following two consecutive inferences regarding continuous P -space:

(i) A deductive one, concluding that *any* basis whatever of nontrivial metrical relations on that space "must come from somewhere else" *in the sense of being constituted by an extrinsic and transported device*. For, as he tells us regarding continuous P -space, "there is requisite some means of carrying forward one magnitude as a measure for the other" via a "superposition of the magnitudes to be compared."

(ii) An *inductive* inference, concluding that there is a dynamical basis for any particular metrical behavior of the extrinsic device under transport, and thereby for the metrical relations generated by it. Note that *only* the *second* of these inferences, which is inductive, pertains to the *dynamical* conception of the metric! Thus, Riemann did not *confine* himself to calling prophetically for a dynamical explanation of the metrical structure of continuous P -space. Instead, his brilliant anticipation of the dynamical status of the space-time (and space) metric in the general theory of relativity was inductively grounded on RMH.

To obviate possible serious terminological confusion, we must distinguish and emphatically dissociate from one another the following two senses of "intrinsic" (and correlatively of "extrinsic"):

(1) The sense of "intrinsic" which is to explicate Riemann's aforementioned notion of "implicit" (German: "*schon enthalten*") as applied to a metric.

(2) The quite different sense of "intrinsic" (German: "*inner*") familiar from Gauss's theory of surfaces that are embedded in Euclidean 3-space. Concerning the *latter* sense, the geometer Eisenhart writes ([21], p. 84): "The geometrical properties of a surface in euclidean 3-space which depend upon the fundamental form [$ds^2 = g_{ik} dx^i dx^k$; $i, k = 1, 2$] alone as distinguished from its properties as a subspace of the enveloping euclidean space are called *intrinsic*. We apply this term to the properties of any [n -dimensional space] V_n depending only upon its fundamental form." Thus, it is said in *this* sense that the constant positive Gaussian curvature of a sphere embedded in Euclidean 3-space is an "intrinsic" property of that surface, whereas its so-called *mean* curvature (cf. [12], pp. 94–95) is not. What is asserted here is that the Gaussian curvature is determined by the *two* dimensional metric ds^2 alone, while the mean curvature depends on properties of the surface which involve its embeddedness in the three-dimensional hyperspace. But it is *not* being asserted that any two-dimensional metric or the geometrical properties engendered by such a 2-metric are "implicit" in that given two-dimensional manifold as Riemann employs that term. In short, intrinsicness in the sense growing out of the Gaussian theory of surfaces is relative to an already given n -dimensional metric on the n -space which is said to possess properties that are

intrinsic in the Gaussian sense. But this sense does not pertain to the status of the n -dimensional metric itself. Hence the assertion of the existence of metrical properties of an n -dimensional space which are intrinsic in this generalized Gaussian sense in no way contradicts Riemann's claim that no such properties are "implicit" in such a space. Thus, it is in this generalized Gaussian sense that Wheeler devotes a section of one of his papers ([72], p. 66) to the "Intrinsic Three-Geometry and Extrinsic Curvature of a Spacelike Hypersurface."

As a final preliminary to our attempted explication of Riemann's notion of "implicit," it will be useful to comment on some ideas of W. K. Clifford, who had published an English translation of Riemann's Inaugural Dissertation in 1873.²¹ In a chapter of Clifford's posthumously published work *The Common Sense of the Exact Sciences* [9], which he had dictated in 1875,²² Clifford said:

The measurement of distance is only possible when we have something, say a yard measure or a piece of tape, which we can carry about and which does not alter its length while it is carried about. The measurement is then effected by holding this thing in the place of the distance to be measured, and observing what part of it coincides with this distance ([9], p. 48). . . . The reader will probably have observed that we have defined length or distance by means of a measure which can be carried about *without changing its length*. But how then is this property of the measure to be tested? We may carry about a yard measure in the form of a stick, to test our tape with; but all we can prove in that way is that the two things are always of the same length when they are in the same place; not that this length is unaltered. . . .

Is it possible, however, that lengths do really change by mere moving about, without our knowing it?

Whoever likes to meditate seriously upon this question will find that it is wholly devoid of meaning ([9], pp. 49–50).

These words of Clifford's, spoken only two years after the publication of his translation of Riemann's Dissertation, seem to me to give a lucid statement of part of the philosophical import of RMH. But in the published abstract of his 1870 lecture "On the Space Theory of Matter," Clifford had declared:

I hold in fact

- (1) That small portions of space *are* in fact of a nature analogous to little hills on a surface which is on the average flat; namely, that the ordinary laws of [Euclidean] geometry are not valid in them.
- (2) That this property of being curved or distorted is continually being passed on from one portion of space to another after the manner of a wave.
- (3) That this variation of the curvature of space is what really happens in that phenomenon which we call the *motion of matter*, whether ponderable or etherial.
- (4) That in the physical world nothing else takes place but this variation, subject (possibly) to the law of continuity ([10], pp. 568–569).²³

Wheeler ([75], pp. 32–34) articulates his own present-day elaboration of Clifford's original conception in the following words:

First, we'll accept Einstein's general relativity or "geometrodynamics" in its standard 1915 form, translated of course into the appropriate quantum version. Second, we accept

²¹ See Clifford's posthumously published [9], pp. xxx and 247.

²² This date is given in Karl Pearson's Preface ([9], p. lxiii).

²³ For references to other, earlier publications where Clifford's 1870 lecture may be found, see [75], p. 41, ref. 3.

as tentative working hypothesis the picture of Clifford and Einstein that particles originate from geometry; that there is no such thing as a particle immersed in geometry, but only a particle built out of geometry. Third, we recognize that we have not come far enough along the road to trace out the consequences of this working hypothesis. We don't know how to describe particles as built out of geometry. Therefore we ask, assuming that particles are built out of geometry, what ideas can we get as to the building process from the very existence of particles.

This conception appears to give rise to an important question, as long as the appropriate quantum version of geometrodynamics does not tamper with the continuity of the Riemannian manifold by discretizing the manifold itself, and does not undermine the very concept of the curvature of space. For Clifford tells us in his third and fourth theses that matter and radiation are indeed *constituted* by curved space. And it would certainly *seem* that *in that case*, the very curvature of space *cannot* itself depend on *any* "non-implicit" metrization by transported bodies. Instead, this curvature would need to obtain with respect to a metric implicit in empty space. We saw earlier at the end of section 1(b) that a Cliffordian theory must individuate its events and space points *without* logical reliance on any prior identity of "ponderable or etherial" particles. And we see now that such a theory presumably also needs to postulate an "implicit" metric! But such a denial of RMH is *inconsistent* with Clifford's cited claim that it is "wholly devoid of meaning" to ask whether the lengths of our familiar measuring standards "do really change by mere moving about, without our knowing it." For the supposition of such an unnoticed change would surely have meaning with respect to the presumed implicit metric that confers a curvature on empty space, *provided* that (1) the underlying conception of the empty space manifold has meaning, and (2) there exists a metric implicit to the latter manifold.

See
Append.
§46

Conversely, if RMH is true and no such metric exists, then it is unclear how Clifford's vision as stated can be implemented within the framework of Riemannian geometry. It is very noteworthy but puzzling in this connection that Wheeler conceives of implementing Clifford's vision *without* the presumption of the existence of any implicit metrics, i.e. without a denial of RMH. For in his joint paper with Marzke ([48], p. 48), he is concerned "to arrive at a way to compare an unknown interval CD between two nearby events anywhere in space-time with the interval AB between two fiducial events." These intervals may both be space-like ([48], p. 56). And it is clear that Wheeler conceives of the metrical comparison of AB and CD as depending on a non-implicit, transported standard. For having presented the method of intercomparison worked out by Marzke, the latter author and Wheeler write:

Now for a central issue! Evaluate the ratio CD/AB by one route of intercomparison. Evaluate it also by another route (Figure 3-10). Compare the two values for the ratio. Will they agree? No discrepancy has ever been found. Therefore, it is reasonable to accept the basic postulate of Riemannian geometry, that *the ratio CD/AB is independent of the choice of route of intercomparison.*

... This postulate is not obvious and, in principle, could even be wrong. For example, Weyl once proposed (and later had to give up) a unified theory of electromagnetism and gravitation in which the Riemann postulate was abandoned. In Weyl's theory, two measuring rods, cut to have identical lengths at a point A in space-time, and carried by different routes to a point C , will *differ* in length when they are brought together ([48], p. 58).

Furthermore, these authors consider "what kind of physics would not be compatible with Riemann's postulate," offer the "Validity of Pauli Principle as Partial Evidence for Riemann's Postulate" ([48], p. 60), but conclude ([48], p. 61) that "It would be desirable to have a more *decisive* experimental argument for the Riemannian postulate." I take Wheeler's concern with the empirical validity of the Riemann postulate of which he speaks as betokening his belief that the metrical comparison of AB and CD depends on a nonimplicit transported standard, as required by RMH. For if he thought that he could appeal to an implicit metric, the question of a possible dependence of the ratio of the measures of the two intervals on the route of intercomparison (path of transport) would not even need to arise! How then is the vision of Clifford to be reconciled with Clifford's and Wheeler's apparent endorsement of RMH?

It may be worth remarking here that unless Clifford's vision can give meaning to the *denial* of RMH for its *empty* space, the known feasibility of specifying a so-called "intrinsic coordinate system" under sufficiently heterogeneous conditions in the GTR does not seem to solve that vision's problem of individuating the points of its empty space. Writers on the GTR "call a coordinate system intrinsic if it can be defined uniquely on the basis of local characteristics of the physical situation" ([4], p. 252). Thus, intrinsic coordinate systems are to be contrasted with coordinate systems that exhibit "the ephemeral aspects of choice" ([4], p. 250). Typically, an intrinsic coordinate system for space-time is furnished by four scalar fields whose construction involves the use of the metric. Thus, insofar as the specification of any intrinsic coordinate system makes essential use of the metric, it can serve to individuate the points in keeping with Clifford's vision only if that vision can meaningfully deny RMH for its empty space. And in the latter case, the points must already be given as individuals in the statement of the metric.

We have, of course, made mention of both four-dimensional space-time (PT -space) and of three-dimensional physical space (P -space). But we have so far deliberately given no consideration to what is understood by a three-dimensional P -space in the context of the relativistic theory of space-time. In the general theory of relativity ("GTR"), attention to what constitutes a P -space is very important for the following reason: The GTR allows for the existence of space-times for which there are three-dimensional spaces P_3 such that (1) one and the same four-dimensional space-time has at least one *flat* three-dimensional P -space associated with it and another one which is *curved*, the same metric standard being used to metrize each of these P -spaces, and moreover (2) the curved P -space is finite while the flat one is infinite ([64], Ch. 1, esp. pp. 28–29). Furthermore, the GTR allows spatial geometries which are time-dependent with respect to a specified metric standard and others which are not in the same space-time, as we shall see in section 7 of 2(c), (*ir*).

Our impending characterization of metrics as intrinsic or extrinsic is, of course, intended to be illuminatingly applicable to the space-time metrics of the GTR and to such spatial and temporal metrics as are allowed by it. But the characterization of metrics as intrinsic or extrinsic can be given without attention to the details of metrizing particular families of space-like hypersurfaces, i.e. simul-

taneity slices, in relativistic space-times. Hence we can safely postpone refinements appropriate to the GTR until we *apply* our concept of an extrinsic (“non-implicit”) metric to the metrics encountered in the GTR or allowed by it.

(ii) *Statement of criteria of intrinsicity and extrinsicity for metrics defined on intervals (“m-metrics”).*

We are now ready to characterize metrics which are defined on *intervals* as intrinsic or extrinsic. Earlier (cf. section 2(a), subsection 3), we had occasion to note that, in certain *physical* contexts, *distance* metrics, which are defined on ordered pairs of points, are *physically* symbiotic with, if not parasitic upon length metrics, which are defined on *intervals*. More generally, in any kind of *P*-space or *T*-space, the physical interest of assigning distance numbers to pairs of points derives from the fact that they are the termini of (one or more) *intervals* of specified lengths or measures! Thus, the equality of length and distance in a singly extended *P*- (or *T*-) manifold on which a length function is also defined should be regarded as a condition necessary for the spatial (temporal) relevance of the mathematical distance function. To assure the physical relevance and interest of possible intrinsic *D*-metrics, we shall rely on this intimate connection between *L*-metrics (or *M*-metrics) and *D*-metrics in subsection (iii) of the present section 2(c) to extend our criteria of intrinsicity and extrinsicity to the case of *D*-metrics.

Since our concept of an intrinsic monadic property of an interval will rely on the notions of being a *non-external* and being a *general* property, we shall deal with these notions first. To begin with, I should emphasize that when I speak of a monadic or polyadic “property”, I do so in the *intensional* sense of that term and *not* in the *extensional* sense of a class! Unless this is borne well in mind, the most important of the definitions which are to follow will appear ill-conceived.

(1) *Given* the elements of a manifold, we shall speak of an entity as being “*internal*” to an interval of the manifold (or as being an “*inside*” entity with respect to the interval), iff the existence of the interval depends on the existence of the entity. Thus every element belonging to an interval $[a, b]$ is internal to $[a, b]$ in this sense, whether $[a, b]$ is an interval of some *P*-manifold or of the arithmetic manifold of the real numbers.

(2) Now, in a given manifold, a monadic *property* *P* is said to be “*external*” to an interval possessing it, iff the obtaining of *P* depends on entities which are *not internal* to the interval.

(3) I shall take it to be antecedently understood what counts as a *general* property, and I explain this notion here only to the extent of pointing out the following: If a property is general, then its constitution does *not* involve particular individuals! Thus, the cardinality of an interval is a general monadic property of the interval. But the disjunctive property of being either singleton $\{a\}$ or singleton $\{b\}$ is *not* a general property of any singleton. Nor is it a general property of an interval in a singly extended continuous *P*-space, for example, to be an interval whose respective end-points have the coordinates 3 or 4 and 5 or 6.

In the preceding three definitions, I have invoked the *dependence* of (i) the existence of one entity on that of another and of (ii) the obtaining of a property on there

being entities of a certain kind. I also appealed to the "non-involvement" of particular individuals in the constitution of a property. The cognate relations of dependence and involvement relevant here seem to be tantamount to the relation of logical implication (entailment) among propositions. Hence I am implicitly confronted here by issues that must be faced by *anyone* who postulates an ontological and epistemic status of attributes which provides an underpinning for logical truth. Thus, for example, in his search for a general characterization of law-like sentences, Hempel—in analogy to Carnap's notion of a purely qualitative *property*—speaks of "a purely qualitative predicate as one whose meaning can be made explicit without reference to any one particular object" (*Aspects of Scientific Explanation*, The Free Press, New York, 1965, p. 270). However, Hempel concludes in the same sentence that this characterization "points to the intended meaning [of the term "purely qualitative predicate"] but does not explicate it precisely, and the problem of an adequate definition of purely qualitative predicates remains open."

Since I make no pretense to offering an ontology here, suffice it to say that whatever the status of attributes and of logical truth, I am prepared to recognize that we deal with them by means of language. And, moreover, I would grant, therefore, that the corresponding linguistic criteria employ a certain amount of "regimentation" in Quine's sense. Yet I hope to be right in believing that the regimentation employed in my account (either tacitly or explicitly) permits this essay to contribute to a clarification of the specific issues to which it is addressed.

(4) In a given manifold, a monadic property P is said to be "*intrinsic*" to an interval possessing it, iff P is a *general* property of the interval and *not external* to it. Thus, for an arbitrary interval of a continuous P -space, the property of being 1 meter long is *not* an intrinsic monadic property. And the monadic property of having end-points whose respective coordinates are 3 and 5, possessed by an interval in a singly extended continuous P -space, fails to be intrinsic, because it is both *non-general* and *external*: the obtaining of this property depends on a name-giver, i.e. on an entity *not* internal to the interval.

(5) In a given manifold, a monadic property P is *extrinsic* to an interval I possessing it, iff P is not intrinsic to I .

(6) In a given manifold, a monadic property P is intrinsic to the intervals possessing it, taken *collectively*, iff it is intrinsic to each interval that possesses it.

(7) In a given manifold, a monadic property P is extrinsic to the intervals possessing it, taken collectively, iff P is not intrinsic to them, taken collectively.

(8) *Mutatis mutandis*, there are corresponding definitions of (i) a polyadic (n -adic, $n > 1$) property P_n being *external* to an n -tuple of intervals related by it, and being *intrinsic* or *extrinsic* to an n -tuple, taken singly, as well as of (ii) the n -adic property P_n being intrinsic or extrinsic to the n -tuples of intervals possessing it, taken collectively.

A symmetric, reflexive and transitive dyadic *property* (in the *intensional* sense!) can qualify as intrinsic in accord with the definition (8) just given. There is no standard term for such an equality property as far as I know: the locution "dyadic equivalence property" is an unfamiliar one, and the term "equivalence relation-*intension*" is clumsy. Hence I shall use the familiar term "equivalence relation"

but with the following explicit understanding: the term "relation" is to be understood in the *intensional* sense of relational *property* and NOT in its customary *extensional* sense of a set of ordered pairs. For on the latter customary construal of the term "relation," it would make no sense to speak of an equivalence relation as being intrinsic rather than extrinsic to its relata. In this connection, the reader is asked to recall the anticipatory *caveat* which I issued near the beginning of the overall **Introduction** regarding the intensional construal of ascriptions of the equivalence relation of congruence. And I shall occasionally remind the reader of my special use of the term "equivalence relation" by putting "*dyadic property*" in parentheses next to it.

(9) A partition Π of the class of intervals of a manifold S will be said to be intrinsic to the intervals of S , iff the partition Π is generated by an equivalence relation that is intrinsic to the intervals of S .

As noted at the beginning of the present subsection, we shall first be concerned with any metrics m defined on *intervals* that qualify as intrinsic, and will characterize D -metrics, which are defined on ordered pairs of points, only thereafter. Hence I shall now speak of any metric m defined on intervals without regard to whether m is an M -metric, an L -metric or an R -metric in the sense of the earlier subsections of our section 2 above. Having preempted the capital letter " M " for the particular kind of metrics given by measure functions, I shall use the small letter " m " to represent any of the specified kinds of metric that is defined on intervals. Thus, M -metrics are merely one particular species of m -metrics.

The term "congruence relation" has traditionally been widely used to refer to any kind of metrical equivalence relation among intervals, such as being of the same length. And the term "congruence class" (i.e. class of equivalence classes of congruent intervals) has had a correspondingly permissive usage. But we saw in section 2(a), subsections 1 and 2 that the more technical use of "congruence" is restricted to a particular kind of metrical equality obtaining *with respect to a D -metric*. I shall therefore continue to forego the temptation of using the familiar term "congruence" and will employ the more clumsy term " m -equality." By the same token, I shall use the predicate " m -equal" rather than "congruent," when necessary.

(10) A metric m will be said to be *intrinsic to the intervals* of a manifold S , iff the partition generated by m -equality is a partition intrinsic to the intervals of S in the sense of (9).

(11) A metric m will be said to be *extrinsic* to the intervals of S , iff it is not intrinsic to them.

Independently of the particular choice of a unit or scale factor, an m -metric effects comparisons among intervals not only in regard to equality and inequality. For in the case of unequal intervals, the metric tells us which of the two is the greater one rather than merely that they are unequal. Thus, we can say that the metric has an equality component and a "more-or-less rank order component" over and above a mere "inequality component." Indeed, the m -metric not only gives us a rank order with respect to more-or-less but also gives us a more subtle differentiation beyond the rank ordering: It tells us by what *factor* or multiple one of two

rank-ordered intervals is related to the other independently of the choice of a unit. Thus, we can say that beyond the mere rank-order component of more-or-less, the metric has a *ratio*-component or *size-factor* component. Although the numerical value of the *difference* between the *m*-values assigned to an ordered pair of intervals does depend on the chosen unit, the ratio of these respective *m*-values is, of course, independent of the choice of a unit. Hence, independently of the chosen unit, we can speak of the equality-cum-inequality component, the rank order (more-or-less) component, and of the ratio-component of an *m*-metric.

Now by virtue of being generated by an intrinsic equivalence relation, the intrinsic partition provides an intrinsic basis for the equality component and for the "mere inequality component" of the metric. But it does not as such also provide an intrinsic basis for the metric's more-or-less component, let alone for its ratio component.

However, whenever there is an intrinsic dyadic ordering property of the intervals, then one can define a dyadic ordering relation among the cells (equivalence classes) of the partition by means of it, and the more-or-less component of the metric also has an intrinsic basis. For example, we shall find later in subsection (iv) of section 2(c) that in a discrete *P*-space with or without end-points, the more-or-less component of the cardinality-based intrinsic *m*-metrics $M(I) = 1/n \cdot \bar{I}$ does have an intrinsic basis. By contrast, this merit is not possessed by Massey's pathological progressive *M*-metric on a discrete *P*-manifold *without* end-points though that metric is also intrinsic to the intervals of that manifold.

Furthermore, Massey has given a definition of the conditions under which the ratio-component of an intrinsic *m*-metric has an intrinsic basis as well. Let *x* and *y* be intervals, *m*(*x*) and *m*(*y*) be the respective numbers assigned to them by the *m*-metric, and let *r* be a real number. Furthermore, let *Rxy* be an *intrinsic* dyadic property. Then his definition states the following: The ratio-component of the intrinsic *m*-metric has an intrinsic foundation, iff

$$(r)(\exists R)(x)(y) \left\{ \frac{m(x)}{m(y)} = r \equiv Rxy \right\}.$$

Once we have defined other species of intrinsicity further below and have characterized intrinsic *D*-metrics as well, corresponding definitions of the appropriate species of intrinsic basis of the ratio-component of a given kind of metric will be available by using Massey's definition *mutatis mutandis*.

Clearly, the ratio-component of an intrinsic metric *may* be devoid of an intrinsic foundation just as the more-or-less component may be.

It is very important to note that the linguistic expression *L* which describes or specifies a metric *m* may obscure the fact that *m* does indeed qualify as intrinsic to the intervals of *S*.

To see this, consider a discrete *P*-space *devoid* of extreme points which has been coordinatized by the assignment of the negative and positive integers such that the coordinates reflect numerically the spatial betweenness relating triplets of points. Let *k* and *l* denote the coordinates of the end-points of an interval *I* of that discrete

P -manifold. And now consider the following two descriptions of an M -metric on the intervals of that manifold:

$$(L_1) \quad M(I) = |k - l| + 1$$

$$(L_2) \quad M(I) = \bar{I}.$$

Although L_1 and L_2 differ intensionally, they denote the same mathematical function from intervals I to measure numbers $M(I)$, i.e. L_1 and L_2 have the same extension. Conceiving of any one metric as a function-in-extension, we can say that L_1 and L_2 describe one and the same metric $M(I)$.

Since equicardinality is an intrinsic equivalence relation by our definitions above, it generates an intrinsic partition in the class of intervals of our discrete P -space S . Hence the metric $M(I)$ described by L_2 is intrinsic to the intervals of S by our definition (10). But the *alternative description* L_1 of the very same metric does not make M 's intrinsicness perspicuous, since it refers to coordinate differences $k - l$ rather than to cardinalities $\bar{I}$.

See
rebuttal
Append
§47 to a
criticism
of the
intrinsic
of $M(I)$.

It so happens that even if we were given only the description L_1 in the case of our particular illustrative example of $M(I)$, it would not be hard to see that $M(I)$ does in fact qualify as intrinsic. Nonetheless, the general moral of the difference between L_1 and L_2 in regard to their epistemic role is the following: *Unbeknownst to everyone*, a familiarly used metric m may be *intrinsic* to the intervals on which it is defined. Indeed, while we are able to certify the intrinsicness of a given metric confidently, we can never be certain that a metric which *appears* to be extrinsic is in fact extrinsic with one important exception to be noted in section 2(c), (iv), subsection 7.

We shall return to the ramifications of this epistemic predicament after distinguishing metrics which are intrinsic to the *intervals* of a manifold from those that are only intrinsic to the manifold *as a whole*. To make this distinction, we give a few definitions analogous to those given earlier and then define the *weaker* property of being a metric intrinsic *to the manifold*.

(12) Given the elements of a manifold S , an entity will be said to be *internal* to the manifold S , iff the existence of the manifold depends on the existence of the entity in the logical sense of the comment on definitions (1)–(3) above.

(13) A monadic property P of an interval of S and an n -adic property P_n of an n -tuple of intervals of S are said to be *external to S* , iff the obtaining of P and P_n respectively depends on entities which are *not internal* to S .

(14) A property P or P_n of intervals of S is *intrinsic to S* , iff the property is a general property of intervals and *not external to S* . And the stated kind of property is *extrinsic to S* , iff it is not intrinsic to S .

(15) A partition Π of the class of intervals of S will be said to be *intrinsic to S* , iff Π is generated by an equivalence relation that is *intrinsic to S* .

Recalling (10), we can now define:

(16) A metric m will be said to be intrinsic to a manifold S , iff the partition generated by m -equality is a partition which is intrinsic to S . A metric which is intrinsic to the intervals of S is also intrinsic to S , though not conversely.

(17) A metric m will be said to be extrinsic to a manifold S , iff it is not intrinsic to S .²⁴

For the sake of brevity, I shall speak of a metric m which is intrinsic to the intervals I of a manifold S as being " I -intrinsic." And I shall call a metric that is intrinsic to S as a whole " S -intrinsic."

Furthermore, let us recall the epistemic predicament in which we may be left by the description of a metric. It is then seen to be desirable to characterize the *epistemic* category to which the *description* of a metric belongs by means of the following terms: If it is correctly discernible from the *description* D of a metric m that m is I -intrinsic or S -intrinsic, then we shall say that D is respectively I -intrinsic or S -intrinsic. The failure of a *description* of m to be either I -intrinsic or S -intrinsic does *not* necessarily preclude that m itself is intrinsic. For this failure of the description *may* be due to attributes of the description rather than to the extrinsicity of the metric described by it. Also, if a description D of m is neither I -intrinsic nor S -intrinsic *and* suggests that m is neither I -intrinsic nor S -intrinsic, we shall say that D is extrinsic. Note that while any metric m must itself be either I -intrinsic or I -extrinsic, and either S -intrinsic or S -extrinsic, its *description* D need *not* necessarily belong to one of the three categories of being extrinsic, I -intrinsic or S -intrinsic, if we are left agnostic by it. And clearly the extrinsicity of a *description* D is a matter of individual mentality with respect to what D suggests to a particular person.

A metric m which is in fact I -extrinsic cannot be given an I -intrinsic description. Nor can an S -extrinsic metric be given an S -intrinsic description. But the aforementioned epistemic predicament makes for the fact that either kind of intrinsic metric can be given an extrinsic description no less than the appropriate kind of intrinsic description.

Hence the extrinsic character of *all* descriptions of m which are ever known can furnish only *inductive* grounds for concluding that m is I -extrinsic or S -extrinsic, except for a special case involving a time-dependent spatial metric to be discussed in section 2(c), (iv), subsection 7.

(iii) *Intrinsicity and extrinsicity of D-metrics*

So far we have defined intrinsicness and extrinsicness for metrics which are defined on intervals. And it therefore remains to give the corresponding definitions for D -metrics, which are defined on ordered pairs of elements.

²⁴ The characterization given by definitions (16) and (17) here supplants the altogether different meaning which I gave to the phrase "extrinsic to the manifold" in my "Reply to Hilary Putnam" ([28], p. 218 and [31], p. 20). There I was concerned to allude to the following fact: though a transported standard (e.g. a meter stick) which is used to metrize the intervals of a continuous 3-dimensional P -space is *noninternal* to the *intervals* of that manifold, it is *immersed* in the 3-manifold itself. To allude to this immersedness, I *incorrectly* equated being immersed in the 3-manifold as a whole with being intrinsic to the latter. I did so by wrongly thinking that the transported standard is not extrinsic to the 3-manifold and that therefore the metrics generated by it are not extrinsic to the 3-manifold as a whole. And thus I was led to say of continuous P -space, T -space and PT -space that congruences are extrinsic to their intervals "but not to these manifolds themselves." Massey ([50], p. 342) ever so rightly called this remark "a puzzling nuance." But the generous interpretation which he then went on to give of its meaning was undeserved.

We wish our characterizations to apply not just to D -metrics defined on such physical manifolds as P -spaces or T -spaces, but also to D -metrics defined on arithmetic manifolds. But it does not lessen the applicability of our impending definitions to arithmetic manifolds that we shall build into them a feature necessary to assure their relevance and *interesting* pertinence to these *physical* manifolds. The feature in question arises from taking cognizance of the role that ordered pairs of points can play as termini of *intervals*. I do not deny that from a purely *abstract* point of view, one can deal with D -metrics on ordered point pairs without concerning oneself with that role. But I have already emphasized at the beginning of section 2(c), subsection (ii) that in any kind of P -space or T -space, the physical interest of assigning distance numbers to ordered pairs of points derives from the fact that these pairs are the termini of (one or more) *intervals* of specified lengths or measures.

Demopoulos [13] has charged that I have shown a gratuitous preoccupation with metrics defined on *intervals* to the neglect of D -metrics in dealing with the possible existence of intrinsic metrics in various kinds of P -space. In view of the physical considerations which I have just adduced, I see no merit in this objection. These considerations are the very ones which justify my allowing for the role of point pairs as termini of intervals in the impending definitions of intrinsic and extrinsic D -metrics. Demopoulos dismisses that role as an irrelevance. But he gives no indication whatever as to what intrinsic physical property of point pairs generates a relation of equidistance in continuous P -space that would justify his claim that the corresponding non-trivial ascriptions of equidistance are devoid of a conventional ingredient.

We now take the ordered pairs of elements of a manifold S as our domain. And mindful of our stated need to take cognizance of the role of pairs of points as termini of intervals, we give the following definition, which would *otherwise* seem contrived, even in a singly extended manifold:

(18) An entity will be said to be "*internal*" to a pair of elements, iff it is internal, in the sense of definition (1) above, to at least one closed interval of which the elements of the pair are the termini.

It might be objected nonetheless that this definition is too permissive on intuitive grounds and that it might serve to qualify too wide a class of D -metrics as intrinsic. I reply that this permissiveness certainly cannot lessen the significance of the *failure* of a manifold to possess the resulting kind of intrinsic D -metrics.

Having stated what is to be understood by an entity internal to a pair of elements, we are able to take over definitions (2)–(9) inclusive *mutatis mutandis*, our domain now being ordered pairs of elements rather than intervals. And hence we can define:

(19) A D -metric is intrinsic to the ordered pairs of elements of a manifold S , iff the partition generated by D -equality is a partition intrinsic to the specified ordered pairs. The corresponding definition of "*extrinsic*" is obvious by analogy to (11) above.

Mutatis mutandis, definition (16) can be used to define a D -metric which is S -intrinsic, and (17) for one which is S -extrinsic. To have an abbreviation for the

stronger intrinsicness defined by (19), we shall speak of a *D*-metric as being "*Op*-intrinsic," where "*Op*" alludes to "ordered pair." Thus *Op*-intrinsicness is the *D*-metrical counterpart of *m*-metrical *I*-intrinsicness.

Our epistemic remarks concerning extrinsic, *I*-intrinsic and *S*-intrinsic *descriptions* of *m*-metrics can also be carried over, *mutatis mutandis*, to *D*-metrics, with *Op*-intrinsic *D*-metrics and their descriptions assuming the role of their *I*-intrinsic counterparts.

(iv) *Intrinsic and extrinsic metrics in several varieties of manifolds*

We can ask two questions:

- (1) Is a given manifold endowed with any *I*-intrinsic or *S*-intrinsic *m*-metrics, or with any *Op*-intrinsic or *S*-intrinsic *D*-metrics?
- (2) If so, are these specified kinds of intrinsic metrics *non-trivial*?

It will be recalled that I formulated necessary conditions for the nontriviality of various metrics by reference to *P*- (or *T*-) manifolds and *PT*-manifolds respectively. But I noted in section 2(b), subsection 2 that this involvement of these particular material types of manifold does not commit me to refrain from similarly justifying the imposition of some or all of these nontriviality conditions for certain purposes in the case of some arithmetic manifolds for example.

Accordingly, let me now invoke the nontriviality conditions governing metrics on *P*- (or *T*-) manifolds to give a well-defined meaning to the second of the two questions above, when addressed to the arithmetic manifolds to be considered in this subsection (iv).

1. The real number continuum. Our criteria of intrinsicness and extrinsicness are material rather than formal ones in the sense that they pertain *not* to abstract manifolds *as such* but to their "concrete" realizations—rather than applying alike to abstract manifolds and to their realizations or models.

Thus, in considering whether there are intrinsic metrics on the real number continuum, I understand the latter to be the particular arithmetic *number* objects whose structure the real number postulates were invented to elucidate. Hence, for our purposes, the real numbers are the *particular* model of these postulates furnished by their *customary* interpretation. And the "material" character of our criteria of intrinsicness and extrinsicness will be illustrated by the following fact: in the case of two manifolds each of which is the same kind of one-dimensional continuum, we shall find one of them (the real number continuum) to be endowed with nontrivial intrinsic metrics, while maintaining (on inductive grounds) in accord with RMH that the other manifold (a continuous one-dimensional *P*-space without extreme points) is devoid of any such metrics.

I shall first use the arithmetic manifold of the *absolute* real numbers, including zero ([40], section 63, pp. 52–53) as my manifold, since I can define different *I*-intrinsic *M*-metrics and different *Op*-intrinsic *D*-metrics on that manifold, all of which are *simple*. Thus, I shall first discuss the issues by reference to the *absolute* real numbers on the basis of simple *M*-metrics. But we shall then see that the corresponding conclusions can be established with respect to the class of all real numbers (*negative*, 0, and *positive*) by replacing one of our two simple *I*-intrinsic

M -metrics and one of our simple Op -intrinsic D -metrics by more complicated ones.

For now, turn to the arithmetic manifold of the absolute real numbers, including zero. If a and b are any two absolute real numbers, let $[a, b]$ denote alike the closed, open and half-open intervals corresponding to a and b . And consider two equivalence relations R^1 and R^2 over the set of closed, half-open or open intervals as follows:

1. $(\alpha)(\beta)[R^1\alpha\beta \equiv (i)(j)(m)(n) [\{\alpha = [i, j] \cdot \beta = [m, n]\} \supset \{|i - j| = |m - n|\}]]$
This equivalence relation R^1 partitions the intervals of the absolute real number continuum such that any two intervals $[i, j]$ and $[m, n]$ for which the absolute differences $|i - j|$ and $|m - n|$ are equal, and only such intervals, will be in the same equivalence class of the partition.
2. Similarly, define the different equivalence relation R^2 by the requirement that $|i^2 - j^2| = |m^2 - n^2|$, i.e. by the demand for equality among the absolute values of the differences between the *squares* of the end-points.

Recalling our *caveat* concerning our intensional use of the term “relation” to mean “relational *property*” or “relation-in-intension,” note that both R^1 and R^2 qualify as *I-intrinsic* equivalence relations by our definitions!

Now consider the following two countably additive M -metrics, defined on the intervals of the absolute reals:

$$M_1[a, b] = |b - a|$$

and

$$M_2[a, b] = |b^2 - a^2| = (b + a) |b - a|$$

Let Π_1 and Π_2 be the partitions which are generated by M_1 -equality and M_2 -equality respectively. Clearly the partitions Π_1 and Π_2 of the domain of intervals into equivalence classes are *different*: Two M_1 -equal intervals are generally *not* M_2 -equal. But precisely these partitions are the ones which are respectively generated by the *I-intrinsic* equivalence relations R^1 and R^2 . And since Π_1 and Π_2 are therefore both *I-intrinsic* by our definition (9), M_1 and M_2 must both be *I-intrinsic* M -metrics by our definition (10).

As anticipated in section 2(c), subsection (ii), we can speak in a clear and obvious sense of our *I-intrinsic* equivalence relation (dyadic *property*!) R^1 as being “the intrinsic basis” of our M_1 -equality and thereby the rudimentary intrinsic basis of the *I-intrinsic* metric M_1 . Here the term “rudimentary” is intended to alert us to the fact that we are talking about the intrinsic basis of the *equality*-component of the metric as distinct from the other components in the sense of section 2(c), subsection (ii). Similarly, R^2 is the different rudimentary intrinsic basis of M_2 . Intrinsic metrics that differ from one another only by a scale factor (nonzero multiplicative constant) clearly have at least *the same rudimentary intrinsic basis*. And if their ordinal or ratio-components have an intrinsic basis, that basis will also be the same for both intrinsic metrics. In our concern with the existence or nonexistence of intrinsic bases for metrics, we shall treat all such only trivially different metrics as one and the same metric. Of course, strictly speaking that is not correct. More generally, for any *I-intrinsic*, *Op-intrinsic*, or *S-intrinsic* metric μ , we shall speak of the

intrinsic basis of μ in the following sense: The appropriately intrinsic equivalence relation which generates the same intrinsic partition as μ -equality will be said to be "the intrinsic basis" of μ . But we shall point out in section 2(c), (iv), 4(β) that the definite article in "the intrinsic basis" would have to be replaced by the indefinite "an" if a given metric possesses two *intensionally* different rudimentary intrinsic bases.

We remarked that it suits our needs here to treat metrics which are only trivially different from one another as one and the same metric. By the same token, we shall now find it useful to speak of the metrizations effected by metrics that are only *trivial variants* of one another as one and the same metrization, as contrasted with being alternative metrizations. For our concern is with the partitions generated by μ -equality rather than with the particular scale-dependent number assigned to any given interval. And hence we shall refer to metrizations as "alternative" only if (and if) they yield *different* partitions into equivalence classes. But it is to be borne in mind that the differing partitions do not preclude our using both of the "alternative" metrizations.

In our stated *strong* sense of "alternative," our two *I*-intrinsic metrics M_1 and M_2 plainly effect alternative metrizations of the absolute reals by issuing in different partitions. Hence alternative metrizability is certainly *not* a sufficient condition for the non-existence of *I*-intrinsic metrics, i.e. for *I*-intrinsic metrical amorphousness. Furthermore, our two intrinsic equivalence relations R^1 and R^2 constitute the respective *I*-intrinsic bases of our two alternative metrics M_1 and M_2 . Therefore, each of our two alternative metrics M_1 and M_2 has an intrinsic basis of its own. It is very important to appreciate the following fact concerning the *I*-intrinsic equivalence relations (dyadic properties!) R^1 and R^2 which constitute the respective intrinsic bases of M_1 and M_2 : R^1 and R^2 *each* obtain as a matter of *absolute real number fact* and *not* as a matter of convention, when they are truly ascribed to two intervals by *M*-equality! Thus we can say that the *pertinent kind of fact* assures the nonconventionality of *each* of these ascriptions. M_1 -equality renders the intrinsic equality of intervals with respect to the absolute difference between their termini, while M_2 -equality renders the likewise intrinsic equality of intervals with respect to the absolute difference between the *squares* of their termini. And the fact that M_1 -equality does not effect the same partition as M_2 -equality clearly cannot detract from the *I*-intrinsicness of either of them. Nor can the nonconventionality of the obtaining of R^1 and of R^2 possibly be impugned by our freedom to choose whether to render one of them rather than the other, both of them, or neither of them by means of an appropriate *M*-equality.

This state of affairs enables us to draw a philosophically important conclusion in anticipation of some of section 3 below: When a certain *I*-intrinsic equivalence relation (dyadic *property*!) R^1 is ascribed by means of M_1 -equality, one cannot gainsay or discredit the nonconventionality of that ascription by pointing out that there also exists another *I*-intrinsic equivalence relation R^2 which issues in a different partition. In other words, it is *not* a necessary condition for the nonconventionality of the ascription of a certain metrical equality that this equality be the *ONLY* intrinsic equality, although it *is* necessary that it be intrinsic. Unless the

metrical equality in question is intrinsic, it will be devoid of a sufficient basis in the *pertinent kind of fact* for *not* being convention-laden. I emphasize the importance of the *pertinent kind of fact*, since an *extrinsic* metrical equality may well have a basis in *some kind of fact or other*. To take a deliberately far-fetched example: the *facts* of *aesthetic* preference and aesthetic equivalence felt by an existing Laplacean demon for the intervals of a manifold might be a basis for a metric on the intervals of the manifold, so that aesthetic equivalence may be a factual basis for extrinsic metric equality on that manifold. But surely such an *irrelevant, extraneous factual basis* would hardly be regarded as establishing the *nonconventionality* of the extrinsic metric equality in question!

Thus, what matters for the nonconventionality of a given *M*-equality is whether the corresponding equivalence relation obtains as a matter of intrinsic fact. As in our examples from the absolute real numbers, there may be more than one *I*-intrinsic equivalence relation. When there is, it is immaterial that we have a choice whether to render one particular intrinsic equivalence relation by the appropriate *M*-equality or another such relation, or both or neither. For the existence of this alternative does not impart any conventionality to such ascriptions of *I*-intrinsic *M*-equality as ensue from exercising options within the framework of the stated alternative. The reader will recall the cognate remarks I made in section 2(b), subsection 2, when discussing our option to tell the spatial story: We can choose to tell or not to tell specified parts of the spatial story metrically, but this option does *not* show that when we do tell it, then the story that is told must be true by convention. A doctor has the choice of telling or withholding his diagnosis from a patient. But the truth or falsity of the diagnosis does *not* thereby become conventional.

Our two nontrivial *I*-intrinsic metrics $M_1[a, b] = |b - a|$ and $M_2[a, b] = |b^2 - a^2|$ on the absolute real numbers have two obvious *I*-intrinsic *L*-metrical counterparts. And they also have two *D*-metrical counterparts which qualify as *Op*-intrinsic, viz. $D_1(a, b) = |b - a|$, and $D_2(a, b) = |b^2 - a^2|$. This result will now be seen to have an important bearing on the range of the beautifully imaginative explication of the intrinsicality of a metric proposed by van Fraassen in his Panel article [68]. Van Fraassen ([68], p. 349 and p. 349, fn. 4) explicitly confined his proposed explication to manifolds which are not only metrically homogeneous but whose elements are of homogeneous constitution in the respective senses set forth in section 1(a) above. The manifold of absolute real numbers is metrically homogeneous with respect to each of the four intrinsic metrics M_1 , M_2 , D_1 and D_2 . But the *elements* of this arithmetic manifold are *not* of homogeneous constitution in the sense of our section 1(a). Hence this arithmetic manifold falls outside the avowed purview of van Fraassen's proposed explication. Yet my characterization of intrinsic metrics pertains to this kind of manifold as well.

Therefore, I should point out by reference to our four intrinsic metrics on the absolute reals what important consequence would follow, if van Fraassen's criteria *were* also to be applied to this arithmetic manifold. In that event, *his* criteria would yield the undesirable result that the absolute reals do not possess any intrinsic metric. To see this, note first that since $D_1(a, b) = M_1[a, b]$ while $D_2(a, b) =$

$M_2[a, b]$, the absolute real numbers admit of distance functions D_1 and D_2 which are *not* constant multiples of one another but which are respectively linear functions of measures M_1 and M_2 . And observe that it then follows from van Fraassen's definitions ([68], pp. 350–351, definitions 1 and 3, and the synonymy on p. 351, fn. 6) that the absolute reals are devoid of any intrinsic metrics. Though I find van Fraassen's explication both ingenious and valuable in its own right, it is seen to be unduly restrictive for my purposes. For we found that the absolute real number manifold is endowed with at least two nontrivial *I*-intrinsic metrics and at least two nontrivial *Op*-intrinsic distance functions.

We used the *simple* functions $|b - a|$ and $|b^2 - a^2|$ on the *absolute* reals to provide simple illustrations of both alternative *I*-intrinsic and alternative *Op*-intrinsic metrizable. If we had attempted to use both $|b - a|$ and $|b^2 - a^2|$ as metrics on the *full-blown* reals (negative, 0, and positive), then we would have encountered the following situation: the function $|b^2 - a^2|$ could then no longer qualify as an *M*-metric, because it would then no longer be additive. Nor could that function still have qualified as a *D*-metric, since it would have yielded the value zero for $D(a, b)$ in the case of different real numbers a and b such that $a = -5$ and $b = +5$, for example.

As anticipated at the start of the present subsection 1 on the reals, we are entitled to do the following: Assert *mutatis mutandis* about the *full-blown* class of reals all the claims of alternative *I*-intrinsic and alternative *Op*-intrinsic metrizable that we made about the *absolute* reals, provided that we replace the simple measure function $M_2[a, b] = |b^2 - a^2|$ and the simple *D*-metric $D_2(a, b) = |b^2 - a^2|$ by suitably different and more complicated ones. That this is indeed the case was pointed out by Massey, who defined the required kind of different countably additive and nontrivial *I*-intrinsic measure function $M_2^*[a, b]$ on the intervals of the full-blown real number continuum. And this function has an *Op*-intrinsic *D*-metrical counterpart on the full-blown reals.

Massey first defines a dyadic relation $\lambda(a, b)$ of *having like signs* among any two reals a and b as follows:

$$\lambda(a, b) \equiv \{(a \geq 0 \cdot b \geq 0) \vee (a < 0 \cdot b < 0)\}.$$

And note that the individual real number zero ingredient in this definition can be eliminated by means of the following description in which no free variables occur:

$$(\exists x)(y) \{x + y = y\}.$$

Then Massey defines the required *I*-intrinsic countably additive nontrivial measure function M_2^* contextually as follows: If $\lambda(a, b)$, then $M_2^*[a, b] = |b^2 - a^2|$, and if $\sim \lambda(a, b)$, then $M_2^*[a, b] = b^2 + a^2$. And if we replace $M_2^*[a, b]$ by $D_2^*(a, b)$, the functions $|b^2 - a^2|$ and $b^2 + a^2$ will do double duty such that D_2^* qualifies as an *Op*-intrinsic *D*-metric.

It is clear, therefore, that our conclusions about the *absolute* reals can be justifiably stated, *mutatis mutandis*, concerning the full-blown reals as well.

When now proceeding to discuss the manifold of rational numbers, we shall be prompted by the same reasons of simplicity to talk about the *absolute* rationals.

including zero, rather than about the full-blown rationals (negative, 0 and positive). But it is to be understood again that this does *not* confine the relevance of our conclusions to the absolute rationals vis-à-vis the full-blown rationals.

2. The absolute rational numbers. The arithmetic manifold of absolute rational numbers, including zero, arranged in the usual order ([40], section 51, pp. 39–40), is a dense denumerable manifold.

Since the absolute rationals are only countably infinite, the following two measure functions on the absolute rational number manifold are only *finitely* additive: $M_1[a, b] = |b - a|$, and $M_2[a, b] = |b^2 - a^2|$. But for reasons quite analogous to those set forth concerning the corresponding countably additive M -metrics on the absolute reals, the present finitely additive M_1 and M_2 are each I -intrinsic and non-trivial.

By the same token, the nontrivial D -metrics $D_1(a, b) = |b - a|$ and $D_2(a, b) = |b^2 - a^2|$ on the absolute rationals are each Op -intrinsic.

Thus, the absolute rationals, like the reals, possess alternative nontrivial intrinsic metrics. And their associated different metric equalities are each nonconventional. Corresponding remarks apply to the manifold of all rationals (negative, 0 and positive). As will emerge from our discussion of a dense denumerable P -space further on, all the I -intrinsic M -metrics on such a P -space that come to mind are *trivial*, in contradistinction to the case of the dense denumerable rationals. And presumably there are no *nontrivial* I -intrinsic, Op -intrinsic or S -intrinsic metrics on such a P -space. This result points up our prior remark, at the beginning of section 2(c), (iv), subsection 1, that our criteria for intrinsicness are “material” rather than formal: Certain important features of the constitution of the rational number elements relevant to the existence of intrinsic metrics on the rationals are *not* adequately captured by characterizing the manifold of the rationals formally as a dense denumerable one.

3. The natural numbers. In the discrete denumerable manifold of the natural numbers, a metric on that manifold specified by $|b - a|$ or by $|b^2 - a^2|$ *cannot* qualify as an M -metric on the intervals of that manifold, since neither function is finitely additive. As will be recalled from the end of section 2(a), subsection 2, the additivity axioms for finitely and countably additive M -metrics make the additivity of the measures of two or more sets conditional on their being disjoint. But it will likewise be remembered that the ordered pairs whose distance numbers are added in the triangle inequality for D -metrics are *not* disjoint but have exactly one common element. Thus, whereas for the manifold of the natural numbers the functions $|b - a|$ and $|b^2 - a^2|$ fail to satisfy the finite additivity requirement for M -metrics, they do satisfy the triangle axiom for D -metrics and indeed both qualify as distance functions.

If we write $D_1(a, b) = |b - a|$, and $D_2(a, b) = |b^2 - a^2|$, we can say that both D_1 and D_2 qualify as nontrivial *Op-intrinsic* D -metrics. And their associated different D -metrical equalities are each nonconventional. If $\bar{I}$ is the cardinality of the interval $[a, b]$, or I , corresponding to the ordered pair (a, b) , we can write

$$D_1(a, b) = \bar{I} - 1,$$

and, since neither a nor b can ever be negative,

$$D_2(a, b) = (b + a) |b - a| = (b + a) (\bar{I} - 1).$$

Evidently D_1 is a linear function of the cardinality of $[a, b]$, but D_2 is a function of both the cardinality and the further variable $(b + a)$. As in the case of the analogous *Op*-intrinsic D -metrics on the absolute reals, this result has a bearing on an important facet of van Fraassen's explication. In commenting on this bearing, let me again emphasize the intended limited scope of van Fraassen's explication to which I called attention in Section 1 on the absolute real numbers.

Van Fraassen refers to the distance function D_1 , which he calls " d ," and to the "natural measure" μ , which is equal to the cardinality. And he writes:

Grünbaum appears to regard d as an intrinsic distance function because it can be defined in terms of the natural measure—or equivalently, in terms of the cardinality of the interval. Definability may here be the crucial notion; it does not seem to me to be too far-fetched to take this to mean that d has to be a linear function of μ , and that if d' is a multiple of d , then d' is a trivial variant of d ([68], p. 350).

But here van Fraassen appended the following precautionary footnote to the word "linear": "The requirement of linearity might perhaps be relaxed in various ways without affecting our conclusions." Our D -metric D_1 shows that an *Op*-intrinsic distance function *can* be a linear function of the cardinality of the corresponding interval. But furthermore, our nontrivially different distance function D_2 demonstrates that a D -metric on the discrete denumerable manifold of the natural numbers can also be *Op*-intrinsic while *not* being a linear function of the cardinality of the corresponding interval. Hence van Fraassen's provision for the relaxation of the requirement of linearity would have to be invoked, if the scope of his explication were to be widened so as to include arithmetic manifolds.

We now take leave of arithmetic manifolds and turn to certain *physical* ones.

To avoid introducing yet another letter designation for a particular species of P - (or T -) manifold, I shall speak of a singly extended manifold having no extreme elements as being "bilaterally boundless." And I shall retain the designation " E -manifold" (or " E -space") for a singly extended manifold having exactly *one* extremity, as well as the locution " F -manifold" (or " F -space") for a singly extended manifold possessing *two* extreme elements. Thus, an E -space is equivalently characterized as unilaterally bounded (or as unilaterally boundless), whereas an F -space is bilaterally bounded.

4. Discrete P - (or T -) manifolds. Our original characterization of a discrete space in section 1(a) was explicitly stated as applying to a *singly* extended manifold. And the metrics we are about to discuss are defined on that kind of P - (or T -) manifold. But I hope that others may elaborate the impending considerations so as to encompass multiply extended granular ("atomic," quantized) P , T , and PT -manifolds.

We shall be concerned with the existence of I -intrinsic, *Op*-intrinsic and S -intrinsic metrics on the following three species of discrete P - (or T -) manifold: Bilaterally boundless, unilaterally bounded, and bilaterally bounded.

Let us recall from section 2(b), subsection 2 our discussion of two nontrivial,

finitely additive M -metrics, both of which are relevant to the first two of these three species of manifold, and one of which is relevant to all three:

- (1) $M_1(I) = 1/n \cdot \bar{I}$, where n is a natural number greater than zero. As explained earlier, we treat all the functions generated by different values of n as one and the same M -metric for our purposes. And we have been calling this one metric "the cardinality-based" measure function.
- (2) Massey's "progressive" M -metric $M_2(I)$, so-called because the *different* measures $1 + s$ which it assigns to distinct singletons are specified by a variable s which ranges over the numbers in the geometric *progression* $1/2, 1/4, 1/8, 1/16 \dots$

The cardinality-based M -metric is relevant to all three of our species of discrete P - (or T -) space, while Massey's progressive one is intended for the two species which are denumerable manifolds.

Let us begin with the *important* species of discrete physical manifold constituted by a denumerable discrete P - (or T -) manifold *without* any extreme element.

(α) Bilaterally boundless discrete denumerable P - (or T -) manifolds

Since equi-cardinality is an I -intrinsic equivalence relation (dyadic property) over the set of intervals, the cardinality-based nontrivial M -metric M_1 is I -intrinsic. But so is Massey's progression metric M_2 . For we saw in section 2(b), subsection 2 that Massey's M_2 -equality generates the same partition as the equivalence relation of (logical) identity. And the latter equivalence relation (dyadic property) is no less I -intrinsic than equi-cardinality. The I -intrinsicity of the "pathological" M_2 is not jeopardized by the fact that M_2 -equality generates a different and indeed much less interesting partition Π_2 than M_1 -equality does: Each cell (equivalence class) of Π_2 contains only a single member, in striking contrast to the partition Π_1 generated by M_1 -equality. And M_2 -inequality goes hand-in-hand with (logical) *nonidentity*.

It is true that M_2 does qualify as an I -intrinsic metric by our definitions (9) and (10) no less than M_1 does. Yet the cardinality-based M_1 does possess a *major distinction* in regard to I -intrinsicity that M_2 clearly lacks. To characterize this important preeminence of M_1 , let us recall our account in section 2(c), subsection (ii) of the "equality component" of a metric, the "more-or-less component" over and above the mere inequality component, and the ratio-component. Then we can see that whereas the equality and inequality components of an I -intrinsic metric will always have an I -intrinsic basis—which we have called the "*rudimentary*" I -intrinsic basis of the metric itself—the "more-or-less component" of such a metric *may* indeed be devoid of an I -intrinsic basis as may the ratio-component. Hence we can say that an I -intrinsic metric M may or may not be "*ordinally I*-intrinsic" over and above being I -intrinsic in regard to M -equality and M -inequality. And we can likewise say that an I -intrinsic metric may or may not also be "ratio I -intrinsic." If an intrinsic metric is indeed ratio I -intrinsic, then it can also be shown to be ordinally I -intrinsic as follows: For any given set of intrinsic

properties, the higher level property of being characterized by at least one member of the given set is itself an intrinsic property. In particular, let Γ be the set of intrinsic dyadic properties which are correlated with ratios $r > 1$ in Massey's definition of ratio-intrinsicality {section 2(c), subsection (ii)}. (An illustration of such a set Γ will soon be given by reference to our cardinality-based M_1 .) Then for any intervals α and β , let $R_\alpha^0\beta$ obtain iff α is related to β by a member of Γ . This dyadic property $R_\alpha^0\beta$ is an *I*-intrinsic basis for the ordinal component of the metric. Thus, ratio *I*-intrinsicality is at least as strong as ordinal *I*-intrinsicality. Indeed ratio *I*-intrinsicality seems to be a stronger property of a metric than ordinal *I*-intrinsicality, although I do not have an example of a nontrivial *I*-intrinsic metric which is *ordinally I*-intrinsic without also being ratio *I*-intrinsic.

In terms of these locutions, it will now be seen that the following turns out to be the case in the context of our bilaterally boundless discrete *P*- (or *T*-) manifold: The *I*-intrinsic cardinality-based metric M_1 is ratio *I*-intrinsic as well, whereas the progression metric M_2 is not even *ordinally I*-intrinsic! In other words, the "ratio component" of M_1 does have an *I*-intrinsic basis, while even M_2 's ordinal component lacks such a basis. Although ordinal *I*-intrinsicality is redundant with ratio *I*-intrinsicality as explained, let us first compare M_1 and M_2 with regard to ordinal *I*-intrinsicality.

We then observe that the dyadic ordering relation of having a higher cardinality (or its converse) is as much an *I*-intrinsic relational property as the equivalence relation of equi-cardinality. Hence there is *also* an *I*-intrinsic basis for saying that an interval belonging to one of the cells of the partition generated by M_1 -equality is greater or smaller than an interval in one of the *other* cells of the partition. This assures the ordinal *I*-intrinsicality of our cardinality-based M_1 . On the other hand, what could possibly be the *I*-intrinsic basis (or even *S*-intrinsic basis) in our *bilaterally boundless* manifold for assigning the greater of two M_2 -measures $1 + s$ to one particular singleton as against another? Or for assigning the higher of two M_2 -measures to one of two distinct intervals each of which has the same cardinality $\bar{I} > 1$? Or yet for making the measures of a pair of singletons be in a certain ratio? Hence we can say that M_1 is distinguished vis-à-vis M_2 at least in regard to ordinal *I*-intrinsicality. By the same token, we can already state a conclusion that will be substantiated in the remainder of this section 2 in conjunction with section 3: By lacking any kind of intrinsic basis (be it *I*-intrinsic or *S*-intrinsic) in the context of our bilaterally boundless manifold, the *ordinal* component of the *I*-intrinsic metric M_2 —and, *a fortiori*, M_2 's ratio-component—is convention-laden (or, if you like, convention—"infested"). On the other hand, M_1 's ordinal component is just as nonconventional as its equality and inequality components. Furthermore, we can now see by reference to Massey's definition of ratio-intrinsicality given in section 2(c), subsection (ii) above that even the ratio-component of our cardinality-based M_1 is *I*-intrinsic.

For simplicity but without loss of generality, let us adopt a choice of unit so that the positive integer n in our cardinality-based $M_1(I) = 1/n \cdot \bar{I}$ is 1. Since the values of M_1 will then all be given by natural numbers, the ratio r of the measures $M_1(x)$

and $M_1(y)$ of any two intervals x and y can then always be written as the ratio of two natural numbers p and q as follows:

$$r = \frac{M_1(x)}{M_1(y)} = \frac{\bar{x}}{\bar{y}} = \frac{p}{q}$$

where $\bar{x}$ and $\bar{y}$ are the respective cardinalities of the intervals x and y . It follows that $q\bar{x} = p\bar{y}$, an equation which can be written in the form of a sum of q terms $\bar{x}$ on the left, and a sum of p terms $\bar{y}$ on the right to make it apparent that we are merely equating certain sums of cardinalities. And we can now say that all and only those intervals x and y whose measures $M_1(x)$ and $M_1(y)$ are in the ratio $r = p/q$ have the following dyadic property $R^{pq}xy$:

$$R^{pq}xy \equiv (q\bar{x} = p\bar{y}).$$

Since it is an *I-intrinsic* dyadic property of intervals that q times the cardinality of the first is equal to p times the cardinality of the second, the set of dyadic properties $R^{pq}xy$ corresponding to all the ordered pairs p, q constitute the *I-intrinsic* basis of the ratio-component of our metric M_1 in the sense of Massey's definition. Thus, M_1 's ratio-component (and thereby its ordinal component) is also *I-intrinsic*.

We saw in section 2(b), subsection 2 that if k is any positive real constant of proportionality in the cardinality-based measure $M_0(I) = k\bar{I}$ on discrete *P*- (or *T*-) space, then the demand that there be a unit interval in that space, i.e. at least one interval I_1 for which $M_0(I_1) = 1$, restricts the values of k to the range $1/n$, where n is an integer. For any interval in the space contains an integral number of points, and hence the intervals I_1 which are chosen as the unit intervals will each contain some integral number n of points. This cardinality n of any unit interval I_1 is an *I-intrinsic* monadic property of I_1 , as is evident from the Russellian definition. The demand that there be at least one interval in the space that receives the measure unity cannot itself be said to have an *I-intrinsic* basis as far as I can see. But we note that any choice of unit made within the confines drawn by that demand can be said to have a monadic *I-intrinsic* basis. In other words, while the restriction of the scale factor k to values $1/n$ does not have an *I-intrinsic* basis, the chosen unit does have such a basis in the sense that intervals which then qualify as unit intervals do so in virtue of the *I-intrinsic* property of cardinality n . Hence it would be possible to speak in this sense of the unit's possession of an *I-intrinsic* basis. By the same token, one could speak of a metric as possibly being *unit I-intrinsic* (or scale factor *I-intrinsic*) along with speaking of its possibly being *ratio I-intrinsic*. But if one thinks of the function of the metric as being that of metrical *comparison*, the unit component might well be left out of account here as relatively unimportant. On this basis, for any given species of intrinsicity to which a given kind of metric belongs, I regard a metric as *maximally intrinsic*, (i.e. maximally *I-intrinsic*, *Op-intrinsic* or *S-intrinsic*) iff the given intrinsic metric is *ratio intrinsic*.

Let us now return to the cardinality-based M_1 and the progression measure M_2 which we had been comparing in regard to ratio *I-intrinsicity*.

It would seem that our M_1 and M_2 are the only *nontrivial I-intrinsic M-metrics* on our bilaterally boundless discrete *P*- (or *T*-) manifolds. If that is so, then we can

conclude from M_1 's superiority over M_2 with respect to ratio I -intrinsicness that the cardinality-based M -metric is *uniquely* distinguished in regard to *maximal I-intrinsicness* from all other nontrivial M -metrics on the given manifold. And if this conclusion is correct, it constitutes a strong vindication, with respect to intervals, of Riemann's previously cited declaration concerning discrete space. Namely, that the metrical "comparison" of "Determinate parts of a manifold . . . as to quantity comes in discrete magnitudes by counting" ([61], p. 413).

Turning to nontrivial Op -intrinsic D -metrics on our bilaterally boundless manifold, we can see that the nontrivial D -metric

$$D_1(a, b) = \bar{I} - 1,$$

where I is the interval $[a, b]$, is Op -intrinsic. For by our definitions (18) and (19), the cardinality $\bar{I}$ of the interval $[a, b]$ of which a and b are the termini is an Op -intrinsic monadic property of the ordered pair (a, b) . And, by the same definitions, the equi-cardinality of two intervals $[c, d]$ and $[e, f]$ is an Op -intrinsic dyadic property of the ordered pairs (c, d) and (e, f) . Moreover, the *ordinal* component of this cardinality-based D -metric likewise has an Op -intrinsic basis for reasons quite analogous to those given for the corresponding virtue of the M -metric $M_1(I) = \bar{I}$.

Note furthermore that $D_1(c, d) = D_1(e, f)$, iff $M_1[c, d] = M_1[e, f]$. Also any two M_1 -equal intervals will likewise be *congruent* with respect to D_1 as defined in section 2(a), subsection 1. And since both M_1 and D_1 are alike cardinality-based, we can say that they have the same intrinsic basis. Thus, we can therefore say of both of them alike that, in Riemann's words, metrical comparison "comes in discrete magnitudes by counting." From this point of view, we need not distinguish them.

It would seem that among nontrivial D -metrics on our bilaterally boundless manifold, D_1 enjoys the same *unique* status of being *maximally intrinsic* as M_1 does among the nontrivial M -metrics on that manifold.

To put this presumed uniqueness of D_1 into bolder relief, let us now explain how we shall introduce an alternative D -metric D_2 . First *coordinatize* the points of our manifold by the assignment of the negative and positive integers and of zero, such that the coordinates reflect numerically the spatial (temporal) betweenness relating triplets of points. Let k and l be the coordinates of points a and b . Note that *unlike* the situation in a discrete space with *one* extreme element discussed in section 2(b), subsection 6, occupying the k^{th} place or the l^{th} place is *not* an S -intrinsic property of the respective ordered pairs of points denoted by (k, k) and (l, l) in our present bilaterally boundless manifold. Instead, occupying the k^{th} place rather than, say, the $(k + 5)^{\text{th}}$ place is solely a matter of our choice of the origin of coordinates. Now recall from subsection 1 on the reals the definition given by Massey of the relational property λ of having like signs in the case of real numbers. Assume the corresponding definition for the corresponding relational property $\lambda_0(k, l)$ in the class of *all* integers (negative, positive, and zero). But be alert to the fact that the integers k and l here are only the respective *coordinates* of the points (instants) a and b of the

bilaterally boundless discrete P - (or T -) manifold which we are about to D -metrize alternatively.

Then we can define a new D -metric D_2 contextually: If $\lambda_0(k, l)$, then $D_2(a, b) = |k^2 - l^2|$, and if $\sim \lambda_0(k, l)$, then $D_2(a, b) = k^2 + l^2$.

It seems safe to say that for *any* particular coordinatization which reflects betweenness, this metric is neither Op -intrinsic nor S -intrinsic. I make this claim of extrinsicity regarding D_2 even though I am mindful of our *caveat* in section 2(c), subsection (ii) not to simply infer the extrinsicity of a metric from the mere appearance of its description. In particular, I do *not* infer that D_2 is extrinsic for each of the specified coordinatizations from the mere fact that D_2 -equality of ordered pairs of points is *not* invariant under different choices of the origin of coordinates (numerical names of the points). For example, a pair of points labeled 0 and 6 respectively will be D_2 -equal to a pair labeled 8 and 10 respectively. But if the origin of coordinates were shifted by one element such that the labels of the first pair become 1,7 and those of the second 9,11, then the two pairs would no longer be D_2 -equal.

The presumed extrinsicity of our present D_2 enables us to comment on a statement which W. Demopoulos made in the belief that it could serve as a criticism of my earlier views on the relation between metrics on *discrete* space and the cardinality of its intervals. He writes: "distance *need* not bear any interesting relation to the cardinality of the intervals with which it is associated. . . . Notice that it is *not* being suggested that cardinality can bear *no* relation to distance, but only that it *need* not" [13]. Note that even for those point pairs whose distance is given by $D_2(a, b) = |k^2 - l^2|$, D_2 *cannot generally* be expressed as $(k+l)(\bar{I} - 1)$, where the cardinality $\bar{I}$ is anyhow only *one* of several variables. That it is not generally permissible in this context to express $|k^2 - l^2|$ as $(k + l) |k - l|$ is evident from the fact that the function $(k + l) (\bar{I} - 1)$ would inadmissibly become negative for a pair of points both of whose coordinates k and l are negative. And *a fortiori* for point pairs whose distance D_2 is given by $k^2 + l^2$, the distance surely does not bear any interesting relation to the cardinality of $[a, b]$.

But the existence of distance functions on our discrete space which are unrelated to the cardinality $\bar{I}$ in the manner of D_2 cannot serve to invalidate either my earlier or my present claims concerning the unique distinction of the cardinality-based D -metric D_1 . For I have never denied that there are distance functions like D_2 . Instead, I have *denied* that the latter are Op -intrinsic while claiming that D_1 *is* Op -intrinsic! In the case of our present bilaterally boundless discrete P - (or T -) space, D_2 not only fails to be Op -intrinsic but likewise does not qualify as S -intrinsic. And we shall soon see that even in a discrete P - (or T -) space with exactly *one* extreme element, the counterpart of D_2 is only S -intrinsic and not Op -intrinsic.

(β) Unilaterally bounded and bilaterally bounded discrete P - (or T -) manifolds

As before, we shall refer to these discrete manifolds more briefly as a discrete E -manifold and a discrete F -manifold respectively. Consider one of these discrete manifolds which is an E -manifold. And suppose we wished to single out the special

ordinal status of its one extreme element metrically by assigning, say, the measure 1.3 to its associated singleton, while assigning the measure 1 to all other singletons of the manifold. Let these assignments be made by a finitely additive measure function which we shall denote by " M_1^* " to distinguish it from the cardinality-based M_1 , which is defined on all three of our species of discrete P - (or T -) space.

As explained preparatory to definition (9) in section 2(c), subsection (ii), we gave an intensional construal of the symmetric, reflexive, and transitive dyadic properties that constitute the rudimentary intrinsic bases of intrinsic metrics. Hence it would appear to be at least conceivable that the partition Π_1^* of the intervals of our E -space effected by M_1^* -equality be generated by two different dyadic properties (intensional "equivalence relations") R^1 and R^2 as follows: R^1 is *only* S -intrinsic (i.e. S -intrinsic but not I -intrinsic), while R^2 qualifies as I -intrinsic. Hence, if we now go ahead and show that Π_1^* is generated by an equivalence relation (dyadic property) R^1 which is *only* S -intrinsic, this does not necessarily show that M_1^* is *only* S -intrinsic as against being I -intrinsic. Indeed, similar considerations suggest that a given metric which is known to be I -intrinsic might conceivably have two intensionally different rudimentary I -intrinsic bases R^1 and R^2 . If there is such a case, it would then be inappropriate to characterize one of them as "*the*" rudimentary I -intrinsic basis in that case.

In the case of the partition Π_1^* effected by M_1^* -equality, there is an *only* S -intrinsic dyadic property R that generates Π_1^* . Recall that the extreme point E_1 of our E -manifold was characterized by means of betweenness in section 1(a) and that betweenness will turn out to qualify as an intrinsic triadic property of points of our manifold in section 4. And let " N_{xy} " represent the dyadic equi-cardinality property of intervals x, y . It is then apparent that Π_1^* is generated by the following property $R\alpha\beta$:

$$(\alpha)(\beta)[R\alpha\beta \equiv \{E_1 \in \alpha \cdot E_1 \in \beta \cdot N\alpha\beta\} \vee (E_1 \notin \alpha \cdot E_1 \notin \beta \cdot N\alpha\beta)]$$

That $R\alpha\beta$ is *only* S -intrinsic rather than I -intrinsic is evident from the following fact: the extreme point E_1 is *not* internal {in the sense of section 2(c), subsection (ii), definition (1)} to intervals α and β that do not contain it, and yet for *such* intervals the obtaining of the relational property R does depend on the existence of E_1 .

Thus M_1^* is S -intrinsic and nontrivial. And although we saw that R 's being *only* S -intrinsic does not guarantee that M_1^* is also *only* S -intrinsic, there is a very strong temptation to suspect it.

Massey has coupled this nontrivial S -intrinsic M -metric on a discrete E -manifold with an example of a correspondingly S -intrinsic M -metric on a discrete F -manifold that singles out the latter's two extreme elements metrically.

Returning to the discrete E -manifold, consider an assignment of the measures M_2 of the finitely additive "*progressive*" I -intrinsic M -metric in the special *ordinal* fashion mentioned in section 2(b), subsection 6: The k^{th} singleton of the E -manifold is to receive the measure $1 + 1/2^k$. Use " M_2^0 " to denote this special *ordinal* version of the progressive measure M_2 appropriate to an E -manifold. Since the k^{th} *point* of this manifold can be characterized in S -intrinsic fashion by means of betweenness, as noted in section 1(a), so also can the k^{th} singleton. But being the k^{th}

singleton is only an S -intrinsic and not an I -intrinsic property of a singleton. Let us apply the term " k -interval" to any interval I that contains the k^{th} point but no preceding point, and denote such an interval by " I_k ." Being a k -interval is on a par with being the k^{th} singleton in regard to being only an S -intrinsic and not an I -intrinsic property. Then we have that for any k ,

$$M_2^0(I_k) = \bar{I}_k + \sum_k^{k+I-1} \frac{1}{2^k}.$$

It will be recalled that the progressive M -metric M_2 placed no restrictions on which particular singleton is to receive a particular measure number $1 + s$. By contrast, in our E -manifold, the ordinal progressive M -metric M_2^0 is able to lay down such a restriction on an S -intrinsic ordinal basis. Accordingly, in the preceding subsection (a), the more-or-less component of the progressive M -metric M_2 was seen to be devoid of an S -intrinsic basis in a bilaterally boundless discrete P - (or T -) manifold. And, by contrast, we now find that in our E -manifold, the more-or-less component of the ordinal progressive M -metric M_2^0 does have an S -intrinsic basis. For being equi-cardinal, having a higher (or lower) cardinality, and being a k -interval are all S -intrinsic properties of intervals. But every interval qualifies as a k -interval for some value of k , and these S -intrinsic attributes are the properties whose possession determines which of two distinct intervals will be assigned the higher of two measures M_2^0 as follows: (1) Just as in the case of the progressive M_2 discussed in section 2(b), subsection 2, no two distinct intervals receive the same measure M_2^0 even if they have the same cardinality, (2) of two distinct equicardinal intervals, the one of lesser k -value as a k -interval will receive the greater measure M_2^0 , and (3) of two nonequi-cardinal intervals, the one of higher cardinality will receive the greater measure M_2^0 .

It therefore seems clear that the measure M_2^0 on our E -manifold is ordinally S -intrinsic. But this virtue of M_2^0 cannot detract from the following presumed unique distinction of the cardinality-based M -metric M_1 , defined on all three of our species of discrete P - (or T -) manifold: M_1 is the only nontrivial M -metric which is *maximally* I -intrinsic in *every* one of these three species of manifold.

Turning to D -metrics on our discrete E -manifold, consider any point a which has the S -intrinsic ordinal property of being the k^{th} point, and also any point b which is the l^{th} -point. Then we can say that an ordered pair of points (a, b) has the S -intrinsic but apparently *non-Op*-intrinsic property of being constituted by the k^{th} and l^{th} points respectively. Now take any two ordered pairs of points α and β such that α is constituted by the i^{th} and j^{th} points while β is constituted by the m^{th} and n^{th} points. Recall from section 2(c), (iv), subsection 1 on the real numbers how we defined the equivalence relations R^1 and R^2 there. In quite analogous fashion, we can here define an S -intrinsic though not *Op*-intrinsic equivalence relation (dyadic property) $R\alpha\beta$ by the requirement that $|i^2 - j^2| = |m^2 - n^2|$. And it appears to be a safe assumption that there is no *Op*-intrinsic dyadic property of the point pairs of our E -space that generates the same partition as this $R\alpha\beta$.

Then we can assert that $R\alpha\beta$ is the S -intrinsic basis of the D -metric

$$D_2(a, b) = |k^2 - l^2|$$

and that $D_2(a, b)$ is only S -intrinsic and not Op -intrinsic. By contrast, the cardinality-based metric $D_1(a, b) = \bar{I} - 1$ is Op -intrinsic in the case of all three of our species of discrete manifold, and indeed maximally so. Besides, we saw in the preceding subsection (α) that the counterpart of D_2 has no presumption of being even S -intrinsic, let alone Op -intrinsic, in the context of the bilaterally boundless discrete manifold.

It emerges therefore that just as the cardinality-based nontrivial measure M_1 was found to be uniquely distinguished as maximally I -intrinsic in every one of our three species of manifold, so also the nontrivial cardinality-based D -metric D_1 is correspondingly distinguished as maximally Op -intrinsic.

5. Continuous P - (or T -) spaces. As remarked in section 2(c), (i), we can postpone refinements appropriate to the status of space and time in Einsteinian spacetimes until we apply our conceptions of intrinsicity and extrinsicity specifically to such relativistic contexts in section 2(c), (iv), subsection 7 below. But for the sake of economy of statement, the generalized formulation of Riemann's Metrical Hypothesis (RMH) to be given in this section anticipates that application by also containing mention of space-time (" PT ") manifolds.

It is to be borne in mind that when we discuss metrics defined on *intervals* of a multiply extended (e.g. three-dimensional) continuous P -space, we always mean *linear* continuous paths {cf. section 1(a)} and *never* the multiply extended kinds of intervals to which area or volume measures are assigned, for example. Thus it will be recalled from section 2(b), subsection 2 that our conditions for the nontriviality of M -metrics pertain to M -metrics defined on singly extended intervals.

We concluded our treatment of criteria of nontriviality for metrics in section 2(b) by pointing out that our concern with the existence of intrinsic metrics on certain kinds of manifolds will focus on nontrivial intrinsic metrics. Our impending statement of RMH will deny the existence of *nontrivial* intrinsic metrics on continuous P - (or T -) manifolds. We should therefore mention first that continuous P - (or T -) spaces (and continuous PT -spaces) do, of course, possess *trivial* intrinsic metrics. That there are at least two such metrics will be evident from the following examples of trivial I -intrinsic M -metrics on these spaces.

Let " A " represent any nondegenerate interval, " $\{a\}$ " any degenerate interval or singleton, and " I " any kind of interval. Recall the details of the conditions governing countably additive M -metrics from section 2(a), subsection 2. Then consider the trivial countably additive M -metric specified by the following pair of conditions: $M(\{a\}) = 0$ and $M(A) = +\infty$. And note that the partition generated by M -equality is an I -intrinsic one, since it is likewise generated by the I -intrinsic dyadic property of equicardinality. Hence the trivial metric M just defined is an I -intrinsic one. In earlier publications (e.g. [24], p. 10), I specifically mentioned metrical equalities based on equicardinality when illustrating Riemann's assertion of the intrinsic metric amorphousness of continuous P - (or T -) space. Clearly, therefore, I meant that a space is intrinsically metrically amorphous iff every *non-trivial* metric on the space is devoid of a rudimentary intrinsic basis, i.e. iff the equality-cum-inequality component of every *non-trivial* metric on the space lacks an intrinsic foundation!

Indeed, unless the requirement of the nontriviality of the metric is built into the very notion of intrinsic metric amorphousness, that notion would be rendered vacuous (inapplicable) by the following fact: Every nonempty space containing intervals on which an M -metric is defined is endowed with at least one trivial I -intrinsic M -metric, viz. $M_0(I) = 0$. For the utterly dull partition generated by this zero metric is also generated by any I -intrinsic *universal* reflexive, symmetric and transitive dyadic property $R\alpha\beta$ exemplified by at least one of the following: $R^1\alpha\beta \equiv \{(\alpha = \alpha) \cdot (\beta = \beta)\}$ or $R^2\alpha\beta \equiv \{(\alpha = \beta) \vee (\alpha \neq \beta)\}$. In particular, the zero metric $M_0(I)$ qualifies as a second example of a *trivial* I -intrinsic countably additive metric on our continuous P - (or T -) manifolds.

In continuation of our preliminary comments on RMH in section 2(c), subsection (i) above, we are now ready to offer a more precise formulation of RMH. On the *empirical* assumption that RMH is *true* in the latter formulation, section 3 will then elaborate its philosophical import in regard to the ontological status of metric ascriptions to intervals or point pairs of continuous P -space, T -space or PT -space. And their ontological status will then be compared with that of the corresponding ascriptions to intervals of a discrete P - (or T -) manifold. It will turn out that a general philosophical moral, only adumbrated in the Introduction, which I shall draw from that comparison would *not* be vitiated even if RMH were in fact false.

Statement of RMH (generalized)

Consider any *continuous* P -manifold, T -manifold or PT -manifold. Call any such manifold a " W -manifold." Let R_1 be any equivalence relation (dyadic property) among intervals of W such that R_1 is S -intrinsic. Recall from definitions (8) and (14) of section 2(c), subsection (ii) that any I -intrinsic dyadic property is also S -intrinsic but not conversely. Furthermore, let R_2 be any equivalence relation (dyadic property) among ordered point pairs of W such that R_2 is S -intrinsic. And recall that Op -intrinsicity entails S -intrinsicity but not conversely.

Now " m_{nt} " will denote any *NON-trivial* metric defined on intervals of W , where the subscript " nt " is intended to remind us of the *crucial* requirement of nontriviality, subject to the following proviso: If W is an Einsteinian space-time and m_{nt} is an indefinite R -metric, then the requirement of being nontrivial is to be understood as calling for the imposition of the *more modest* necessary conditions stated for this important special case in section 2(b), subsection 5. Similarly, " D_{nt} " represents any *NON-trivial* D -metric.

With these *crucial* understandings, RMH (in generalized form) asserts the following conjunction:

Given any W -manifold and any m_{nt} and any D_{nt} , there exists no R_1 such that the S -intrinsic partition of intervals generated by R_1 is ALSO generated by m_{nt} -equality, and there exists no R_2 such that the S -intrinsic partition of ordered point pairs generated by R_2 is LIKEWISE generated by D_{nt} -equality.

Put more concisely, RMH asserts that in any W -manifold, *every* m_{nt} and *every* D_{nt} is devoid of even a rudimentary S -intrinsic basis and hence is *S-extrinsic*. In other words, RMH claims that every W -manifold is S -intrinsically metrically amorphous. As emphasized above, in the latter formulation, the predicate " S -

intrinsically metrically amorphous" must be taken as pertaining only to *nontrivial* *S*-intrinsic metrics, lest the attribution made by that formulation be turned into an egregious and transparent falsehood.

We can pinpoint more precisely the presumed *nonexistence* of a rudimentary *S*-intrinsic basis for m_{nt} -metrics. To do so, we first define what we mean by "quasi-disjoint" nondegenerate intervals. Recalling the definition of an "*I*-triplet" of intervals from section 2(b), subsection 2, we define: Two nondegenerate intervals are "quasi-disjoint", iff neither one includes the other and they are not members of an *I*-triplet. Note that whereas all disjoint intervals are quasi-disjoint, the converse is not true, since there are *quasi*-disjoint intervals which are not disjoint. Consider the partition Π_{nt} of the class of intervals generated by the metric equality of any metric m_{nt} . Take any two quasi-disjoint intervals I_1 and I_2 belonging to the same cell of Π_{nt} as well as any two quasi-disjoint intervals I_3 and I_4 belonging to different cells of Π_{nt} . Then we can say that RMH's denial of the existence of an *S*-intrinsic R_1 that would generate Π_{nt} is rooted more specifically in the following presumed fact: there is no such dyadic property R_1 whose obtaining among intervals I_1 and I_2 would be the *S*-intrinsic basis of their membership in the same cell *and* whose nonobtaining among intervals I_3 and I_4 would be the *S*-intrinsic basis of their belonging to different cells.

That this statement by reference to quasi-disjoint intervals does furnish the greater specificity which I have claimed for it is evident from the fact that if we confined ourselves to intervals which are not quasi-disjoint and to singletons, intrinsic dyadic properties would be available for putting them into different cells of the partition or into the same cell. For example, in the case of two closed intervals one of which properly includes the other, the *I*-intrinsic dyadic property of proper inclusion such that the difference set includes a nondegenerate interval could serve as the intrinsic basis for putting two intervals into different cells of the partition.

It follows from RMH that every nontrivial *m* or *D* time-metric on a continuous *T*-space is extrinsic. RMH also entails that in any *W*-manifold W_n of two or more dimensions, all Riemannian metrics $ds^2 = g_{ik} dx^i dx^k$ ($i, k = 1, 2, \dots, n$) are extrinsic, since they are all appropriately nontrivial. But the entire metric geometry of W_n is determined by the quantities in the metric. Therefore, the metric geometry rests on an extrinsic basis in *our* Riemannian sense of that term, though *not*, of course, in the Gaussian sense which we explicitly distinguished from it in section 2(c), subsection (i)!

We note at once that in its primary formulation given here, RMH is a conjunction of *denials* of *existential* statements! As previously noted, my primary concern will, of course, be with the philosophical moral that can be drawn, *if* RMH is true as a matter of mathematical-cum-empirical fact. Nonetheless, some comments on the evidential support that can be mustered in its behalf are now in order. But it is to be understood that I am not attempting to canvass in adequate detail the full extent to which the empirical RMH can draw support from pertinent results of mathematical topology.

For brevity, let us first introduce the notion of a *nontrivial* equivalence relation (dyadic property) R_{nt} , be it intrinsic or extrinsic. If Π_{nt} is the partition generated

by the metrical equality of some m_{nt} or D_{nt} , and R is a dyadic property which also generates Π_{nt} , then we shall say that R is “nontrivial” and will denote it by the subscript “*nt*.” By the same token, we shall speak of trivial partitions Π_t and trivial equivalence relations R_t corresponding to the metric equalities of trivial metrics m_t and D_t . Using this notation, we can say that according to RMH, any S -intrinsic R that constitutes a rudimentary basis of a metric m or D on any W must be an R_t .

Turning to the support that can be adduced for RMH, let us look at several kinds of S -intrinsic R that constitute rudimentary bases of metrics m with a view to determining whether the members of any of these species must each be an R_t .

For any given W , we can divide the specified class of S -intrinsic R into those which are topological (species A) and those, if any, which are nontopological (species B, which conceivably *may* be empty). And we can, in turn, divide the topological species A into two subspecies as follows: those S -intrinsic reflexive, symmetric and transitive dyadic properties which obtain in virtue of the fact that two intervals possess the same S -intrinsic *monadic* topological property (subspecies A1), and those “purely” dyadic ones which are *not* constituted by the possession of the same monadic S -intrinsic property (subspecies A2).

Topological properties of a point-set of a space have been discussed in the literature by reference to two different classes of homeomorphisms: (1) all homeomorphisms of a topological space onto *itself*, and (2) all homeomorphisms that relate a given geometrical figure F to *other* figures F' . Thus, in the first of these two construals, it is intended that a monadic property ϕ of a set s of a topological space S ([46], p. 66) is topological, iff for every homeomorphism ψ ([46], p. 100) of S onto *itself*, the image of s under ψ has the property ϕ . In this case, we can say that ϕ is a topological property in the *automorphic* sense of the term. But in the second construal, any monadic property ϕ of the figure F would be topological, iff that property is also possessed by *every* figure F' into which F may be transformed by a homeomorphism. In this second case, we can speak of ϕ as being a topological property in the *figure-to-figure* sense of the term. In the illustrations which we are about to give, it will be clear whether one or both of these senses are pertinent.

I shall assume that the cardinality of a continuous interval is the *next* higher one beyond $\aleph_0$. This assumption enables me to classify the particular cardinality of a continuous interval of any given W as an S -intrinsic *monadic* property, no less than the properties of being an infinite set or of having a denumerable subset. Thus, given this assumption, the *equicardinality* of two intervals of W is an S -intrinsic dyadic property belonging to subspecies A1. But it would constitute no difficulty for my purposes, if it were not the case that a continuous interval has the next higher cardinality beyond $\aleph_0$. For in the latter eventuality, the equicardinality of such intervals would belong instead to the different subspecies A2 of dyadic properties while being no less S -intrinsic.

For a given W , return to our topological subspecies A1 of S -intrinsic dyadic properties R , and consider the equicardinality instance $R^1\alpha\beta \equiv (\bar{\alpha} = \bar{\beta})$. Evidently $R^1\alpha\beta$ is an S -intrinsic topological property of type A1 in both the automorphic and figure-to-figure senses. And we already saw early in the present subsection 5 that

$R^1\alpha\beta$ is an R_t , as required by RMH. This observation applies to any of our W manifolds.

For our next example, consider the intervals of a W_n constituted by the two-dimensional Euclidean topological plane of points, where each point is represented by an ordered pair (x, y) of real numbers. For any S -intrinsic monadic topological property P possessed by each of two intervals α and β of this W_n , consider the dyadic property (equivalence relation) $R\alpha\beta \equiv \{P(\alpha) \cdot P(\beta)\}$. All of the instances of $R\alpha\beta$ generated by the various monadic properties P belong to our subspecies A1. And we can now see that in our chosen W_n , all such $R\alpha\beta$ are instances of R_t as required by RMH.

Take the automorphic type of homeomorphism specified by the conjunction $x \leftrightarrow x$ and $y \leftrightarrow 2y$. Furthermore, take an instance of the kind of closed interval on the y -axis that can be represented by $[0, y]$ as well as a closed interval $[x_1, x_2]$ on the x -axis, which is denoted by I_x . Now suppose that $[0, y]$ and I_x each possess some S -intrinsic monadic topological property P and hence possess the dyadic property $R\alpha\beta$. But then the specified automorphic homeomorphism will carry the closed interval $[0, y]$ into the closed interval $[0, 2y]$, which properly includes it. Consequently $[0, 2y]$ will possess the property P no less than $[0, y]$ and I_x . But this means that all three of these closed intervals belong to the same cell of the partition generated by $R\alpha\beta$ even though one of them properly includes another. By our section 2(b), subsection 2, this partition must therefore be a *trivial* one, and $R\alpha\beta$ must be an instance of R_t , as claimed by RMH.

I shall content myself with these examples from our subspecies A1 without making any attempt to deal with subspecies A2, let alone with the problematic species B.

We have looked at examples of S -intrinsic rudimentary bases R and found them to be trivial ones R_t . Let us note further that for any given W , at least to my awareness, there is no *known* R_{nt} that qualifies as an S -intrinsic rudimentary basis of any m_{nt} or D_{nt} .

For example, in a W_n constituted by a 3-dimensional P -space, consider the m_{nt} -equality defined by having the same length in meters. This m_{nt} -equality *can*, of course, be based on the following somewhat clumsy property: Permitting the application of the same number and/or of equal parts of meter sticks, or of a given meter stick in part and/or whole the same number of times. Clearly, the latter dyadic property is not only nontrivial but also S -extrinsic. But the question is: Can we claim to know that there exists no *other* dyadic property R_{nt} which is S -intrinsic and yet *also* generates the same partition Π_{nt} as the meter stick metric m_{nt} ? We must ask this question. For if there were, Π_{nt} and thereby the meter stick's m_{nt} would qualify as S -intrinsic by our definitions (15) and (16). If the m_{nt} furnished by meter sticks is *not* time-dependent—as we are assuming until we get to subsection 7 below—then we certainly *cannot* claim to know deductively that the meter stick metric is an S -extrinsic one. Such a conclusion certainly does not follow deductively from the mere fact that an identifiably S -extrinsic R_{nt} generates the same partition as m_{nt} -equality does. And we recall the *caveat* of our section 2(c), subsection (ii) to the effect that it is quite possible for an intrinsic partition Π_{nt} to be generated by an

m -equality whose metric m is *described* extrinsically. Hence in our present case of a time-independent spatial metric it is quite possible that there exists some S -intrinsic R_{nt} which generates the same partition as having the same length in meters. This is so, although the description of the metric constituted by length-in-meters does involve the meter stick, an entity which is *not internal* to the manifold S , as defined in our (12) of section 2(c), subsection (ii).

But while we have no deductive proof in our present time-independent case that the meter stick metric is S -extrinsic, we certainly do not have any reason to think that it is S -intrinsic. And the extrinsicity of the *description* of the meter stick metric inductively creates some presumption that this metric is S -extrinsic.

Corresponding remarks apply much more generally to any time-independent spatial Riemannian metric $ds^2 = g_{ik} dx^i dx^k$ ($i, k = 1, 2, 3$) on our 3-dimensional P -space. Here as well, we cannot infer its S -extrinsicity *deductively* from the fact that coordinate names are used in its description. But in the absence of any known intrinsic basis for any such metric ds , the latter's extrinsic appearance inductively permits the presumption that any such metric ds is S -extrinsic.

So much for inductive support for RMH.

6. Dense denumerable P - (or T -) spaces. We saw in section 2(b), subsection 2, that a nontrivial measure, whether finitely or countably additive, must assign the measure *zero* to every singleton $\{a\}$ of a dense P - (or T -) manifold. But the necessary condition $M(\{a\}) = 0$, is, of course, not sufficient for the nontriviality of a measure function on a dense P - (or T -) manifold. For we saw in the preceding subsection 5, that in the case of *continuous* space, this condition is satisfied by the following two *nonetheless trivial* M -metrics M and M_0 on that space: (1) $M(\{a\}) = 0$; $M(A) = +\infty$, where A is any nondegenerate interval, and (2) $M_0(I) = 0$.

Thus, when we inquired in subsection 5 whether there are any *intrinsic* metrics on continuous P - (or T -) manifolds, we found that both of these two metrics M and M_0 are I -intrinsic but trivial.

When now inquiring into the existence of intrinsic metrics on dense *denumerable* P - (or T -) manifolds, be they trivial or nontrivial, let us *not* confine our consideration of M -metrics to those for which the measure of a singleton is zero. First consider a measure function, to be denoted by " K ", which assigns the same *positive* finite real number k to all singletons: $K(A) = +\infty$, and $K(\{a\}) = k$, where all nondegenerate intervals A have the same cardinality $\aleph_0$. In contrast to continuous space, a countable union of singletons of our present denumerable dense space *can* form a nondegenerate interval. Bearing this in mind, we note that the measure K is *countably* additive, and, of course, trivial. Furthermore, since the partition which K -equality generates in the class of intervals is generated by the I -intrinsic dyadic property of equicardinality, K is an I -intrinsic countably additive M -metric but trivial.

Now examine another M -metric N on our dense denumerable kind of space as follows: $N(\{a\}) = 0$ and $N(A) = +\infty$. In assigning the *same* non-zero value to each interval A , we were careful *not* to assign a *finite* constant nonzero value. For such an assignment would have precluded the finite additivity of the resulting function and therefore would have disqualified it as an M -metric. This trivial metric

N generates precisely the same partition as equi-cardinality and hence as the metric K . Therefore, the trivial N is I -intrinsic, just as the trivial K . But since a countable union of singletons of our present denumerable dense space can form a non-degenerate interval, the metric N is only *finitely* additive, whereas K was found to be *countably* additive. Observe that K and N differ only in the respect that K assigns a positive value to all singletons while N assigns zero to them. Observe furthermore that the metric M on dense *nondenumerable* P - (or T -) space makes the very same assignments to intervals $\{a\}$ and A as N does for dense denumerable space. But whereas M is countably additive in its nondenumerable space, N is only finitely additive in its denumerable space. On the other hand, the trivial zero metric $M_0(I) = 0$, which generates the same partition as the I -intrinsic universal equivalence relation discussed in subsection 5, is countably additive and I -intrinsic in denumerable dense space no less than in continuous space.

Not all of the trivial I -intrinsic measure functions on a dense denumerable P - (or T -) manifold are quite as uninteresting as the ones we have mentioned. For there is a modified form Q of Massey's progression measure specified by the following countably additive assignment of numbers to the denumerable infinity of singletons $\{a\}_k$:

$$Q(\{a\}_k) = \frac{1}{2^{2k-1}}.$$

The measure function Q assigns a *finite* measure to every interval which is *different* for every interval. It is I -intrinsic since the dyadic property of logical identity is its rudimentary I -intrinsic basis. And it is trivial, because it violates the second of our nontriviality conditions.

Let us temporarily *disregard* questions of *intrinsicity* of M -metrics. And let us attend to the above findings concerning the additivity properties of certain M -metrics on our two kinds of dense P - (or T -) space for the following purpose: To state their bearing on the status of Zeno's metrical paradox of extension ([26], Ch. III and [27], Ch. III) which P. W. Bridgman has expressed in the following words:

With regard to the paradoxes of Zeno . . . if I literally thought of a line as consisting of an assemblage of points of zero length and of an interval of time as the sum of moments without duration, paradox would then present itself. ([7], p. 490).

Let us now make several comments on the bearing which our results pertaining to the M -metrics M_0 , M , and M_0 , K , N on nondenumerable and denumerable dense space respectively have on Zeno's metrical paradox of extension.

(1) Dense *denumerable* P - (or T -) space.

No M -metric Δ which requires that $\Delta(\{a\}) = 0$ and that $\Delta(A) \neq 0$ can be *countably* additive. And the only countably additive Δ which assigns zero to all singletons is the trivial zero metric M_0 . Indeed, we saw in section 2(b), subsection 2 that there are no *nontrivial* countably additive M -metrics on our dense denumerable space, although there are such finitely additive ones.

Now suppose that someone were to require on physical grounds that Δ be non-trivial and hence, in any dense space, fulfill *at least* the conditions $\Delta(\{a\}) = 0$ and $\Delta(A) \neq 0$. Also suppose that, *again on physical grounds*, someone were to require further that a measure function on any P - (or T -) space containing infinitely many elements be countably additive. According to at least one mathematically literate classics scholar ([47], p. 66) and to classically literate mathematicians ([35], p. 11), Zeno himself did indeed impose this requirement of countable additivity. And Zeno was able to draw the following correct conclusions: (i) A countably additive M -metric Δ *cannot* consistently satisfy the requirements $\Delta(\{a\}) = 0$ and $\Delta(A) \neq 0$ in any dense denumerable space, and (ii) in a dense denumerable space, a countably additive Δ cannot consistently satisfy both the requirement $\Delta(\{a\}) = k$, where k is a positive constant, *and* the requirement that for at least some A , $\Delta(A) \neq +\infty$. Zeno had imposed the requirement of countable additivity on physical grounds. And he was then able to argue that any physical theory which requires either that $\Delta(\{a\}) = 0$, $\Delta(A) \neq 0$ or that $\Delta(\{a\}) = k$, $(\exists A)[\Delta(A) \neq +\infty]$ must deny that P - (or T -) space is denumerably dense.

Obviously, *finitely* additive M -metrics on dense denumerable space can *consistently* satisfy the pair of requirements $\Delta(\{a\}) = 0$, $\Delta(A) \neq 0$. And I have always been well aware of the truism that *finitely* additive measure functions on dense denumerable space do *not* run afoul of Zeno's paradox of extension. But when I discussed the status of dense denumerable space vis-à-vis Zeno's paradox of extension, I did so in the following context: The measure function Δ was required to be *countably additive* on physical grounds, and on the same grounds Δ was required to satisfy the pair of minimal nontriviality conditions $\Delta(\{a\}) = 0$, $\Delta(A) \neq 0$. Hence it was a bit ungenerous of Hilary Putnam to write that "Grünbaum has asserted that Zeno's paradox *depends for its solution* on countable additivity" ([56], p. 222n; italics are Putnam's).

(2) Dense *nondenumerable* (continuous) P - (or T -) space.

By contrast to the situation in denumerable dense space, a countably additive M -metric Δ on a continuous space does *not* run afoul of Zeno's paradox of extension upon the imposition of the requirements $\Delta(\{a\}) = 0$, $\Delta(A) \neq 0$. Incidentally, in this context the definition of linear "interval" given in Section 1(a) can be liberalized ([42], p. 200, and [50], p. 343) so as to allow the full Euclidean topological straight line and either one of two of its "halves" (on either side of a given point) to qualify as an interval.

This brief digression from our concerns with *intrinsic* metrics on dense denumerable P - (or T -) space does bear out a point that I have sought to make in earlier writings on Zeno's paradox of extension. To wit, *given* the imposition—on physical grounds—of the requirement of *countable* additivity of M -metrics, the denumerable and continuous species of dense P - (or T -) space differ importantly as follows: The requirements $\Delta(\{a\}) = 0$, $\Delta(A) \neq 0$ do generate a Zenonian inconsistency in the denumerable case, whereas they do not do so in the continuous case. And if the conjunction of the three requirements is held to be justified on physical grounds

along with the denseness of P - (or T -) space, then this result is an adequate logical reason for rejecting the denumerable species of dense space in favor of the continuous variety.

Let us return to our present main topic of intrinsic metrics on dense denumerable space. In the context of the rather vague adumbration of the concepts of intrinsicity and extrinsicity in my earlier writings, some of my earlier claims about the status of intrinsic metrics on dense denumerable space were plausible. But we shall now see that on the more precise construal of these concepts presented in the present essay, certain earlier claims of mine need to be corrected or clarified more precisely than I have done heretofore.

I had noted ([24], pp.10–11) that “since the distinction between denumerable and super-denumerable dense sets was almost certainly unknown to Riemann, it is likely that by ‘continuous’ he merely intended the property which we now call ‘dense’.” Evidence of such an earlier usage of ‘continuous’ is found as late as 1914.” Hence it can be presumed that RMH can be asserted with respect to dense P , T or PT manifolds rather than merely with respect to continuous ones. And I shall indeed assume RMH to be true in this liberalized form, as I anticipated in my 1968 “Reply to Putnam” ([28], pp. 237–238 and [31], pp. 35–36). Then we can see that the following 1963 statement of mine calls for correction and/or clarification over and above those I gave in 1968 ([28], [31]), and some of those given in Massey’s Panel article ([50], pp. 335–336):

But one cannot invoke densely ordered, *denumerable* sets of points (instants) in an endeavor to show that discontinuous sets of such elements may likewise lack an intrinsic metric . . . even without measure theory ordinary analytic geometry allows the deduction that the length of a *denumerably* infinite point set is intrinsically zero ([24], p. 10).

In terms of the notions presented in this essay, this statement calls for several comments.

(1) If the intrinsic metrics at issue are *trivial* ones, then dense denumerable space is indeed endowed with them, as illustrated by our countably additive M -metrics M_0 and K as well as by our finitely additive one N . But then so is continuous space endowed with *trivial* intrinsic metrics, as illustrated by our countably additive trivial M and M_0 on that space. It is true that although N and M each generate the equicardinality partition in their *respective* spaces, N is only finitely additive while M is countably additive. But the context in which I made my cited statement quite incorrectly links this difference between the additivity properties of N and M to a difference between dense denumerable and continuous space in regard to the status of trivial intrinsic metrics.

(2) If the intrinsic metrics at issue are *nontrivial* ones and if RMH is asserted for denumerable dense space as well, then it is literally false to say, as I did in the cited 1963 statement, that “one cannot invoke densely ordered, *denumerable* sets of points (instants) in an endeavor to show that discontinuous sets of such elements may likewise lack an intrinsic metric” ([24], p. 10). This much I already acknowledged in 1968 ([28], p. 237–238 and [31], pp. 35–36).

(3) Even if the term “length” in the cited statement were charitably read to mean

M -metrical measure—a reading which is undeserved as shown by the statement—it would be very misleading to say that “the measure of a *denumerably* infinite point set is intrinsically zero.” What should be said instead is that the trivial zero metric $M_0(I) = 0$, which is the only countably additive M -metric that assigns zero to all singletons, is an *I-intrinsic* metric.

7. Time-dependent spatial metrics and the dynamical conception of the space-time metric.

In the present subsection, I wish to exhibit the relevance of our notions of extrinsicality and intrinsicity to the following several items: (1) time-dependent spatial metrics encountered in the general theory of relativity (“GTR”), (2) the GTR’s dynamical conception of the space-time metric as well as its dynamical conception of such metrics of space or time as depend on the metric tensor of space-time. And the reader can then form a judgment on the merits of Earman’s contention that “Grünbaum’s thesis of the metrical amorphousness of space and time does not illuminate and does not draw support from the GTR” [16]. Section 3 below will develop the import of the present subsection in regard to the presence of conventional ingredients in the time-dependent spatial geometries encountered in the GTR and in its metricized space-time.

Mindful of our preliminary remarks at the end of section 2(c) (i), it will be well to give a very brief account of certain pertinent facets of relativistic space-time, and of the status of a three-dimensional P -space in the context of a relativistic space-time (PT -space). Before doing so, let us mention a few matters concerning Newton’s P -spaces and Newtonian time.

In addition to Newton’s *absolute* three-dimensional P -space, there were—as he explains in the Scholium of his *Principia* ([55], pp. 6–12)—different “relative” three-dimensional P -spaces in motion with respect to absolute P -space and with respect to one another. Thus these different relative three-dimensional P -spaces were exemplified by different secondary inertial frames in uniform translatory motion with respect to the primary inertial frame of absolute space. But Newton emphasized that, in his view, the relative 3-spaces were endowed with the same *unique* intrinsic spatial equalities as absolute space. For “true,” “real,” “absolute” motion which is uniform covers “truly” (intrinsically) equal intervals of absolute space in truly (intrinsically) equal time intervals, and “Absolute and relative space are the same in figure and magnitude” ([55], p. 6). If we *wish* to speak of the space-time of Newtonian physics, we can say that its global topology was presumed to be that of the Euclidean 4-plane.

In section 1(a) and section 2(b), 2, we noted that in treating all points of a P -space as being of *homogeneous* constitution, one *abstracts* from those diverse *extra*-geometrical (physical) entities and attributes which serve to individuate them for almost all P -spaces. (There we mentioned discrete P -space with one extreme element as an exceptional case with respect to the requirement of an *extra*-geometric basis for individuation.) It is such abstracting which permits the treatment of the points of P -space as distinct and yet homogeneous given individuals. Similarly

for the elements of space-time. Thus J. L. Anderson writes in his recent treatise on relativity:

The common feature of all space-time theories is the use of the concept of the *space-time point*. It corresponds to where and when a physical event, for example, the collision of two mass points, takes place. The collection of all such points, that is, the collection of the wheres and whens of all possible physical events, constitutes *space-time*. The distinguishing feature of a particular point of this space-time is that it has no distinguishing features; all points of space-time are assumed to be equivalent ([2], p. 4).

For brevity, I shall also follow the usage of some relativists and shall refer to a space-time point as a "*world point*" to convey that it is an element of the four-dimensional *PT*-manifold in contradistinction to the *points* of a *P*-manifold and the instants of a *T*-manifold. Mindful of the foreign, external character of the individuating devices in the case of the homogeneous elements of almost all these important manifolds, I quite generally took the elements of our manifolds as *given* when defining an entity as "internal" to a manifold in definition (12) of section 2(c), (ii), just as in definition (1) there. But I also emphasized in section 1(a) that precisely because the physical devices that serve to individuate the points of a continuous three-dimensional *P*-space and the world-points of space-time are extrinsic to these manifolds, the so-called space-theory of matter seems to contain a vicious circle in its principle of individuation. It is to be understood hereafter that when a particular event, like the collision of two particular mass points, is used to specify a particular world point, the latter world point, qua entity of the *PT*-manifold, is *not* endowed with all of the features of the physical event that serves to identify it. Nonetheless, I shall occasionally employ the rather common parlance of also referring to an element of a *PT*-manifold as a punctal "event."

Unlike Newton's theory, the GTR countenances continuous space-times that do not possess the global topology of the Euclidean 4-plane. And some of its space-times may in principle have a feature which we shall now describe by reference to a three-dimensional manifold F_3 of non-intersecting time-like world-lines. Such a manifold F_3 can be specified by the nonintersecting careers of a suitable continuum of distinct particles. The feature in question is that in some of the GTR's space-time manifolds, there may be *no* family F_3 such that *every* world-point ("event") of the space-time belongs to *exactly one* member of the specified family of time-like world lines ([2], Ch. I and [43], p. 74). We confine our attention to the case of those regions of those space-times that do have the following property: In the region of the space-time in question, there exist families of time-like world lines such that for a given three-dimensional family F_3 , each world-point of space-time in the region belongs to exactly one member of F_3 .

Having introduced this restriction, we can then do both of the following: (1) In a given space-time, we can specify the system of spatial points constituting a particular three-dimensional *P*-manifold P_3^0 by means of a particular family F_3^0 of non-intersecting time-like *world-lines*, and (2) we can claim that all of the events of the appropriate region of space-time occur at a *unique* spatial point of P_3^0 . Any three-dimensional *P*-manifold P_3 specified in this way can then serve as a spatial reference

frame. And at least one ordered triplet of real numbers can be assigned to each of its spatial points as coordinates ([43], p. 74 and [54], pp. 234–236).

Thus, speaking of the much more special case of a Galilean frame of reference in Minkowskian space-time, J. L. Synge writes:

When we speak of *space* in Newtonian physics, we think of a 3-dimensional manifold of points. When we speak of space-time in relativity, we think of a 4-dimensional manifold of events. The link between the two lies in the fact that the set of parallel [time-like] world lines we have been considering form a 3-dimensional manifold of world lines, and it is these *world lines* that are to be identified with the *points* of familiar space. The observer is justified in speaking of "space" as a Newtonian physicist might speak of it, provided that he admits two things. First, a "point" of his "space" is really a world line in space-time, and, secondly, that it is *his* "space" and not a universal "space", since the directions in space-time of the parallel world lines may be chosen arbitrarily ([66], pp. 51–52; italics in the original).

Observe that the time-like world lines which correspond to the points of a P_3 -space were postulated to *fill* the given region of space-time. For this reason, it would be wrong to view the specified three-dimensional manifold of world-LINES as a proper submanifold of the four-dimensional manifold of world-POINTS constituting the given region of space-time. More precisely, since the *spatial* points (elements) of a P_3 -manifold generated from a given space-time are *not* also *world* points (elements) of that space-time but are correlated with certain *sets* of world points, the three-dimensionality of P_3 does *not* permit the inference that P_3 is a proper submanifold of the four-dimensional space-time. But, of course, a three-dimensional space-like hypersurface in space-time *is* such a proper submanifold.

We see that in the GTR, no less than in Newton's physics, we can specify the points of a three-dimensional P -space by a suitable system of nonintersecting *time-like world-LINES*, each of these world-lines being taken in its entirety in the appropriate region of space-time. Thus, as stressed by Synge, an infinite time-like (equivalence) *class* of event-elements of space-time (e.g. a class of genidentical event-elements of space-time) corresponds to a single element of the P_3 in question. And the very nature of the correspondence which generates a P_3 from space-time shows the following: Intrinsically, the points of any P_3 -space and the P_3 -manifold constituted by them are time-INDEPENDENT entities! In other words, with respect to all of its *intrinsic* properties, any existing P_3 -manifold, qua manifold of spatial points, is a timeless entity (in the sense of definition (4) of section 4 below).

This result will now enable us to establish the extrinsicality of a time-dependent space-metric employed in cosmology. The space points of the spatial reference frame which is metrized by this metric are held to be specified by ideally ubiquitous punctal galaxies. And the spatial reference frame in question is employed by isotropic and homogeneous models based on the so-called "cosmological principle" ([5], pp. 11–15 and 65–67). I refer here to the coordinates and the space-metric associated with the Robertson-Walker line element ([1], Ch. 11, and [5], pp. 129 ff.), when the latter is expressed in a particular coordinate system. Let $d\tau$ be the four-dimensional separation, r , θ , and ϕ the space coordinates, t the proper time read by standard clocks stationed at the galactic space points, K a constant which may be

0, 1 or -1 , and let $R(t)$ be a function of the proper time t . Then the Robertson-Walker line element can be written as

$$d\tau^2 = dt^2 - R^2(t) \frac{dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2}{\left(1 + \frac{K}{4} r^2\right)^2}$$

Let us take a nonconstant function $R(t)$ which is allowed by the field equations of the GTR (with or without cosmological constant) for the case of the homogeneous isotropic models to which this space-time metric $d\tau$ pertains. And note that the entire spatial part of $d\tau$ contains the time-dependent quantity $R^2(t)$ as a common factor.

A few words concerning the general connection between the infinitesimal space-time metric $d\tau$ and the infinitesimal spatial metric $d\sigma$ of a particular spatial reference frame in the GTR need to be said now before we consider the spatial metric $d\sigma$ generated by the Robertson-Walker $d\tau$ for the class of spatial intervals of the spatial reference frame coordinatized by r , θ and ϕ . Let Latin indices take the values 1, 2, 3, 4 and Greek indices range over 1, 2, 3 with summation over the appropriate range of values in the case of a repeated index in accord with the usual summation convention. In any tetrad of coordinates x^i ($i = 1, 2, 3, 4$) used below, x^4 will be the time-coordinate while x^α ($\alpha = 1, 2, 3$) will be the space-coordinates. Thus, given a four-dimensional coordinatization $\{x^i\}$, the world lines of the space points of the corresponding reference frame will be given by $x^\alpha = \text{constant}$ ([54], pp. 234–236). Hence we represent the invariant space-time separation by

$$d\tau^2 = g_{ik} dx^i dx^k,$$

where the g_{ik} are the components of the space-time metric tensor appropriate to the coordinates $\{x^i\}$. Then the infinitesimal spatial distance $d\sigma$ measured in the corresponding spatial reference frame between any two neighboring points $A(x^\alpha)$ and $B(x^\alpha + dx^\alpha)$ is given ([54], p. 238) by

$$d\sigma^2 = |\gamma_{\alpha\beta} dx^\alpha dx^\beta|, \quad \text{where} \quad \gamma_{\alpha\beta} \equiv g_{\alpha\beta} - \frac{g_{\alpha 4} g_{\beta 4}}{g_{44}}.$$

We see that the spatial metric coefficients $\gamma_{\alpha\beta}$, which determine the spatial geometry of the reference frame, are *not* equal to the spatial part $g_{\alpha\beta}$ of the four-dimensional metric tensor g_{ik} , unless $g_{\alpha 4} = 0$ ($\alpha = 1, 2, 3$). Since a coordinate system for which $g_{\alpha 4} = 0$ at all world points is said to be “time-orthogonal,” we can say that the time-orthogonality of the coordinate system is a necessary and sufficient condition for writing the spatial metric as $d\sigma^2 = |g_{\alpha\beta} dx^\alpha dx^\beta|$. But note that the right-hand side of this equation is obtained from setting the time coordinate increment dx^4 equal to zero in the four-dimensional expression for $d\tau^2$. Hence the time-orthogonality of the coordinate system is a necessary and sufficient condition for obtaining the infinitesimal spatial distance $d\sigma$ between points $A(x^\alpha)$ and $B(x^\alpha + dx^\alpha)$ by setting dx^4 equal to zero in the right-hand side of $d\tau$ and dropping minus signs.

Returning to the Robertson-Walker (hereafter “R-W”) line element, we see that the coordinate system r , θ , ϕ , t is time-orthogonal, since no terms $g_{\alpha 4}$ appear

in the expression for the invariant four-dimensional $d\tau$ in terms of these coordinates. Hence we can obtain the spatial metric $d\sigma^2$ for the R-W reference frame by setting $d\tau^2 = 0$ in the expression for $d\tau^2$ (after discarding the minus sign). Thus, we have

$$d\sigma^2 = R^2(t) \frac{dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2}{\left(1 + \frac{K}{4} r^2\right)^2},$$

where $R^2(t)$ was assumed to be a nonconstant function of the proper time t .

This space-metric is clearly a time-dependent one, but it has certain special features. To state these, let us consider intervals of the R-W *spatial* reference frame, each such spatial interval being specified by an appropriate bundle of nonintersecting time-like world lines. Then we note the following features of the time-dependent space-metric $d\sigma$: (1) Any two space intervals AB and CD which have equal lengths $d\sigma$ at some one time t_1 will be $d\sigma$ -equal at any other time t , but (2) the length $d\sigma_1$ of a given space interval AB at some fixed time t_1 will generally *not* be equal to the length $d\sigma$ of that *self-same* space interval at other times t . The latter feature is hardly surprising, since this metric is to have the capability of characterizing the expansion of the universe. Indeed the field equations of the GTR (with or without cosmological constant) allow solutions in which $R(t)$ *never* has the same value at any two different times, so that *no* space interval AB will *ever* have the same length $d\sigma$ at *any two different times*.²⁵ We shall now concern ourselves with the latter type of R-W space metric $d\sigma$ in considering whether this metric is intrinsic or extrinsic.

To do so, consider the relation of $d\sigma$ -equality and the character of any equivalence relation (dyadic property) $R_{\alpha\beta}$ of *space intervals* that might be the basis of $d\sigma$ -equality. Then we find that the relation of having the same length $d\sigma$ qualifies as an equivalence relation in the class of R-W space intervals at any one time but *only* at any one time! For while any given interval AB does stand in the relation of having the same $d\sigma$ -length to itself at any one time, AB *never* stands in that relation to itself at two distinct times! But no *S-intrinsic*, let alone *I-intrinsic* reflexive, symmetric and transitive dyadic property of *space intervals* could constitute the basis of this kind of time-dependent $d\sigma$ -equality. For we saw that any and all of the *intrinsic* properties of a spatial manifold P_3 and of any of its intervals obtain *timelessly*! And thus any *intrinsic* reflexive (as well as symmetric and transitive) dyadic property of *spatial* intervals would obtain between a given interval at one time and itself at any other time. In other words, no equivalence relation (dyadic property) among

²⁵ One such solution, obtained from the equations on certain assumptions as to the matter-energy distribution and with vanishing cosmological constant Λ , features a Euclidean ($K = 0$) three-dimensional P_3 -space (though *not* a flat, Minkowskian space-time), yielding $R = At^{2/3}$, where A is a constant. Another solution (de Sitter model), obtained via $\Lambda \neq 0$, features $K = 0$ and yields $R = R_0 e^{t/T}$, where both R_0 and T are constants.

The value $K = 0$ corresponds to Euclidean P_3 -space for the following reason. If we put $e^{-g(t)} \equiv 1/R^2(t)$, then the *spatially constant* but *time-dependent* Gaussian curvature of the P_3 of our reference frame is given by the product $Ke^{-g(t)}$. K can be 0, 1 or -1 , and $e^{-g(t)}$ is, of course, always positive. Thus the three possible values of K respectively yield Euclidean, spherical and hyperbolic spatial geometries of spatially constant curvature. And for a given matter-energy distribution, the solution of the field equations will determine both K and $R(t)$.

spatial intervals that fails to obtain between an interval and itself at *different* times can be an *S*-intrinsic let alone an *I*-intrinsic dyadic property. Hence no *S*-intrinsic equivalence relation can be the basis of R-W $d\sigma$ -equality, where $d\sigma$ is construed as a metric on the intervals of the *spatial* manifold P_3 . The latter $d\sigma$ -equality is a dyadic property similar to the one of being of the same age (in microseconds) in the class of humans. By contrast, in a discrete *P*-space, no *M*-equality based on the *I*-intrinsic equivalence relation of equicardinality could be time-dependent.

It follows that the time-dependent R-W *space* metric $d\sigma$ is *S*-extrinsic. To obviate a possible misunderstanding, let me stress that my proof of the extrinsicity of $d\sigma$ is claimed to hold with respect to the *spatial* manifold P_3 when $d\sigma$ is construed as a metric on the spatial intervals of that manifold P_3 . If $d\sigma$ were construed instead as a metric on the space-LIKE intervals of the four-dimensional R-W space-time, then I would need to invoke *RMH* in order to justify the assertion that, *thus* construed, $d\sigma$ is extrinsic to space-time. By contrast, as Massey has noted, there is at least one *granular space-time* to which a time-dependent metric on its "space-like" intervals is intrinsic. But there may be other granular space-times having the property that no *such* metric is intrinsic to them.

We can now state how to view the time-dependent $d\sigma$ -equality as an equivalence relation. If α and β are space intervals, while t_1 and t_2 are instants of time, consider the two ordered pairs (α, t_1) and (β, t_2) . Then these two ordered pairs can be said to stand in the equivalence relation of $d\sigma$ -equality, iff the length of α at t_1 equals the length of β at t_2 . Indeed, to take fuller cognizance of the role of time in $d\sigma$, we should view it in a way that I have deliberately postponed discussing until after establishing its extrinsicity *in its construal as a metric on the intervals of P_3* . Specifically, in the four-dimensional space-time of the R-W line element $d\tau$, consider the family of three-dimensional proper submanifolds of world-points generated by the slices $t = \text{constant}$ of the coordinate system in which we expressed the R-W $d\tau$. Each such slice of constant t is a space-like hypersurface H .²⁶ And the members of the one-parameter family $H(t)$ of these space-like hypersurfaces clearly do not intersect each other. This one-dimensional family of hypersurfaces specifies a one-dimensional *T*-manifold, since there is a 1-1 correspondence between the members of $H(t)$, which are simultaneity classes, and the elements (instants) of *T*.

Viewing any given space interval AB of the R-W reference frame at different times t_1 and t_2 then takes the form of considering the following two *disjoint* intervals of world points in the given space-time: the space-time interval I_1 specified by the intersection of $H(t_1)$ with the bundle AB of time-like world-lines which are the careers of the spatial points of the space interval AB , and the space-time interval I_2 formed by the intersection of the different hypersurface $H(t_2)$ with the same bundle AB . Let us now employ a kind of cosmological super meter stick which yields the values of $d\sigma$, and which is an entity *external* to both the R-W P_3 -manifold and the R-W space-time manifold in the sense of our definitions (12) and (1) of section 2(c), (ii). The hypothetical continuum of particles constituting the cosmic

²⁶ For a definition of "hypersurface," see [70], p. 18, and for a statement of the conditions under which a hypersurface is space-like, see [2], p. 155.

meter stick each have time-like world lines of their own, and each such world line is labeled by the appropriate graduation of that galactic meter stick. Then the assignment of lengths $d\sigma$ to the two space-like intervals I_1 and I_2 —which are different slices through one and the same bundle of time-like world lines AB —might take the following form.

We bring a particular pair of the meter stick's graduated time-like world lines into intersection with the pair of world-lines of the space points A and B such that the two world-points of the intersections belong to $H(t_1)$ ([66], pp. 31–32). This process yields the length $d\tau$ of the space-like space-time interval I_1 via the simultaneous coincidence at t_1 of two points on the meter stick with the two space points A and B . For note that since the R–W coordinate system is time-orthogonal, the spatial length $d\sigma_1$ of the *space*-interval AB at t_1 is *equal* (to within a factor of $\sqrt{-1}$) to the space-time “length” $d\tau$ of the space-time interval I_1 . We carry out a similar application of the meter stick at the later time t_2 . Given our special assumptions concerning the function $R(t)$ for an expanding universe, we shall find that the space interval AB coincides with a larger portion of the meter stick at t_2 , yielding a length $d\sigma_2$ for AB at t_2 which is greater than $d\sigma_1$. And the length $d\sigma_2$ is equal (to within a factor of $\sqrt{-1}$) to the space-time length $d\tau$ of the space-time interval I_2 . We chose a cosmic meter stick as the external metric standard for graphic simplicity. But any one of a family of concordant, interchangeable kinds of external metric standards could have been used to obtain the same $d\sigma$ -equalities and inequalities. And what matters is that it is the relation of the space interval AB to an external metric standard that changes with time, *not* the *I*-intrinsic or *S*-intrinsic properties of AB *qua* interval of our R–W reference frame P_3 .

The difference between the purely spatial and the space-time way of viewing the space interval AB makes itself clearly felt in the following way: The space-time intervals I_1 and I_2 are *disjoint*, although they are each generated by a time slice through one and the same bundle of world-lines AB . But clearly the space interval AB is identical rather than disjoint with itself *qua* interval of the R–W reference frame P_3 . We can think of the time-independent space interval AB as being obtained by appropriately stringing together, as it were, a continuum of certain coordinatized space-like intervals of space-time, like our I_1 and I_2 , taking exactly one such interval from each member of the family $H(t)$. And we can say for brevity that AB is *not* “self-congruent” in time iff a space interval AB has a time-dependent infinitesimal length $d\sigma$. (I employed the latter locution in [28], Ch. II).

Our proof of the *S*-extrinsicality of the time-dependent $d\sigma$ pertaining to our *spatial* R–W reference frame P_3 holds quite generally for *any* time-dependent space metric $d\sigma = \sqrt{\gamma_{\alpha\beta}} dx^\alpha dx^\beta$ whose assignments of equal and unequal lengths to spatial intervals of a P_3 at different times rest on a common basis of comparison. In any time-dependent space metric $d\sigma$, at least one space interval will *not* be self-congruent in time. For if all of the space intervals of a P_3 were self-congruent in time, then each one would have the same length at all times. And in that case the space metric could not be time-dependent. But since at least one space-interval will not be self-congruent in time if $d\sigma$ is time-dependent, the reasoning we used to

establish the S -extrinsicity of the time-dependent R-W $d\sigma$ to P_3 applies quite generally to any time-dependent $d\sigma$, when construed as a spatial metric on the intervals of a P_3 .

This important result, to which we alluded at the very end of section 2(c), (ii), has two very significant consequences which I now wish to articulate in turn.

1. In the case of a time-dependent space metric $d\sigma$, we do *not* need to invoke RMH to certify the S -extrinsicity of $d\sigma$ to P_3 : Even those who reject, doubt, or are uncommitted regarding RMH must admit that the time-dependent space metrics $d\sigma$ are S -extrinsic to P_3 . And in section 3 we shall develop the import of the extrinsicity of a metric for its convention-ladenness.

2. The S -extrinsicity of a time-dependent $d\sigma$ *cannot* be impugned by the following twofold fact: (a) the GTR prescribes a *unique* metric $d\sigma$ (to within a choice of unit) in any given spatial reference frame of a given space-time, and (b) the partition of space intervals effected by $d\sigma$ -equality *at any one time* is *unique*. The complete compatibility of the latter twofold kind of uniqueness with the extrinsicity of $d\sigma$ needs to be borne in mind for our section 3 below. For there it will serve to invalidate Earman's belief that he is posing a difficulty for my account of the significance of time-dependent $d\sigma$ when he writes:

The congruence classes [partitions] of infinitesimal spatial intervals will be different at different times, but . . . at no instant is there any freedom in the choice of the measure of these intervals in the chosen frame [apart from the unit used] [16].

In my 1968 book ([28], pp. 208–209, and 231), I did not discuss our cosmological R–W example of a time-dependent $d\sigma$ but used the more earthy illustration of a disk (to which a third space axis was attached) rotating with variable angular velocity ω with respect to an inertial frame in Minkowski space-time. Earman [16] denies that the latter constitutes a meaningful illustration of a time-dependent spatial geometry. Indeed, unless the space coordinates of the space points of the rotating disk frame can be associated with a time coordinate x^4 so as to yield a time-orthogonal coordinate system, Earman denies the legitimacy of speaking of a spatial $d\sigma$ -geometry of the disk frame, even if the disk rotates with constant ω . I now wish to consider the issues raised by this claim before going on to the remaining matters of the present subsection. Hereafter, I shall speak of the values of the spatial distance and of the space-time separation as being “the same” or “equal” even when they are related by $\sqrt{-1}$ or by some other non-zero constant factor. And whenever the space-time separation of two events is imaginary, I shall drop the factor $\sqrt{-1}$ and will take only the real multiplier to be its value.

Let x^4 be a coordinate such that the equation $x^4 = \text{constant}$ specifies a family $H(x^4)$ of nonintersecting space-like hypersurfaces in the Minkowski space-time of our rotating disk frame K . Suppose further that the coordinates of the fixed space points of the disk frame K and our x^4 coordinates are *not* a time-orthogonal system, i.e. that the components $g_{\alpha 4}$ ($\alpha = 1, 2, 3$) of the space-time metric tensor g_{ik} generally do *not* vanish. Now consider a pair (a, b) of events both lying within a *single* space-like hypersurface $x^4 = \text{constant}$. As we saw earlier, when time-

orthogonality *fails* to obtain, the value of $d\sigma$ assigned in the frame by the function $d\sigma = \sqrt{\gamma_{\alpha\beta}} dx^\alpha dx^\beta$ as the spatial distance between *such* a pair (a , b) of events is generally NOT equal to the *four*-dimensional separation $\sqrt{g_{\alpha\beta}} dx^\alpha dx^\beta$ of a and b . Earman claims that whenever it is generally *not* the case that $d\sigma = \sqrt{g_{\alpha\beta}} dx^\alpha dx^\beta$, then $d\sigma$ *fails* to qualify as a spatial metric for the space-like hypersurfaces $H(x^4)$ and for the given spatial frame except locally. And the argument he gives for this contention is as follows: the correspondence rule linking the GTR to the special theory of relativity ("STR"), i.e. the requirement of the local validity of the STR demands that the infinitesimal *spatial* distances $d\sigma$ between world points of a space-like hypersurface in the given spatial reference frame P_3 be *equal* to their four-dimensional separation. Speaking of a spatial reference frame K , Earman writes:

According to one of the basic correspondence postulates linking the GTR to the special theory of relativity (hereafter, STR), the spatial distance $d\sigma$, as measured in K , from a point P_1 of K to a point P_2 of K infinitesimally distant from P_1 is at a given moment equal to the distance as measured in a local inertial frame K_0 which is at rest with respect to P_1 at the given moment. It follows that $d\sigma$ is the length of a space-time displacement which is orthogonal to the world line of P_1 and which lies between the world lines of P_1 and P_2 [16].

To see whether Earman's argument here can serve to establish his requirement of time-orthogonality, picture a set of rods at rest on a uniformly rotating disk such that these rods are each *unit* rods after suitable corrections have been made to allow for deformations due to centrifugal forces, etc. as explained by Møller ([54], p. 223). Let a unit rod be available at any spatial point P_n of the disk which is to be used as a base point for the determination of the spatial distance to a neighboring spatial point. Then the local validity of the STR requires the following: For each world point on the world line of any spatial point P_n , it is possible to introduce locally Minkowskian coordinate systems in which, at the given world point, the g_{ik} have the values η_{ik} and the *first* partial derivatives of the g_{ik} vanish. Thus, at each space-point P_n of the disk and for each world-point on the world-line of P_n , there will exist a pair of events E_n and E_0 at the two ends of the rod which are *simultaneous* in the particular co-moving local inertial frame I_n , and therefore the *four*-dimensional separation of *that particular pair of events* E_n , E_0 at the ends of the unit rod will, of course, be unity. But if the space-like hypersurfaces $H(x^4)$ are so chosen—as indeed they are in the treatments by Møller ([54], pp. 226 and 240–241) and by Reichenbach ([59], section 44, pp. 172–178)—that the pair of events E_n and E_0 do *not both* belong to any member of $H(x^4)$, then we can do the following: Combine the event E_n at the P_n -end of the rod with an event E_1 at the other end of the rod such that E_1 is DIFFERENT from E_0 , while E_n and E_1 *both* belong to a member of $H(x^4)$. And now I can collect from the endpoints of each of the rods associated with the various P_n a pair of events E_n , E_1 such that all of the following conditions are met: (1) The individual pairs E_n , E_1 are *not* simultaneous in their respective local co-moving inertial systems I_n , but they all belong to one of my $H(x^4)$, (2) the space-time separation $d\tau$ of each pair E_n , E_1 is *less than* 1, but for each pair the space coordinates x^α on the disk are such that the function $d\sigma = \sqrt{\gamma_{\alpha\beta}} dx^\alpha dx^\beta$ does have the value 1, and the inequality $d\tau \neq d\sigma$ is, of course, due to the fact that the coordinate system $\{x^\alpha, x^4\}$ is not time-orthogonal ($g_{\alpha 4} \neq 0$).

How then is it incompatible with the stated local validity of the STR to do the following: To take $d\sigma = 1$ to be the *spatial* distance on the disk frame K between E_n and E_1 for every such pair, even though their four-dimensional separation is less than 1? It seems to me a *non sequitur* for Earman to declare that such measurements $d\sigma$ "can be made only at the expense of abandoning the correspondence postulates which link the GTR to the STR." If there is no time-orthogonal coordinate system that can be associated with the spatial points of the rotating disk in Minkowski space-time, then Earman's insistence on time-orthogonality has the following physically stultifying consequence: We cannot speak of the spatial geometry of the rotating disk, as Einstein did in section 3 of the fundamental 1916 GTR paper, where he argues that this geometry is hyperbolic ([19], p. 116). In elaboration of Einstein's compact argument there, the presentation by Møller ([54], sections 84 and 90), which I followed in [28], pp. 200–209, does the following. If T is the time coordinate of the global inertial frame I ("ground system") with respect to which the rotating disk has angular velocity ω , the x^4 coordinate associated with the disk frame K is taken to be a coordinate time t given by $t = T$. Hence one and the same family $H(x^4)$ of space-like hypersurfaces is associated with the disk observer's space coordinates x^a and with the ground system's different space coordinates X^a . But whereas the K -frame's coordinates $\{x^a, t\}$ are not time-orthogonal, the I -frame's coordinates $\{X^a, T\}$, where $t = T$, are.

This difference in orthogonality can be made very intuitive as follows. Disregarding the third spatial dimension of both K and I , we consider the circular projection C in I over which the adjacent disk is rotating. Then we can picture the world lines of the space points X^a in C as a family of concentric right circular cylinders *perpendicular* to C whose axis of symmetry is the world line of the space point which is the common center of C and of the disk. In this picture, our family $H(x^4)$ of hypersurfaces is depicted by horizontal circular sections parallel to C and perpendicular to the last cylinder's surface. What of the world line of a space point P_n of the disk frame K whose radial coordinate in K is r ? If the disk rotates with uniform angular velocity ω , the world line of P_n will be a *helix* on the cylinder whose radius is r , and if ω is not constant, the world line of P_n will be a spiral on that cylinder. Clearly, neither the helix nor the more general spiral is orthogonal to the horizontal circular sections which depict the hypersurfaces $t = T = \text{constant}$.

By using rigid rods at rest in I to metrize each member of $H(x^4)$, the I -frame observer is using the FOUR-dimensional space-time metric to metrize each of them and, of course, obtains a Euclidean spatial geometry on each of them. By using (corrected) rigid rods at rest on K to metrize each member of the *identical* family $H(x^4)$, the disk observer is NOT employing the space-time 4-metric and obtains a hyperbolic spatial geometry on each of them. *For each member of $H(x^4)$, it is a matter of alternatively metrizing one and the same space-like hypersurface.* It is explicit in both Møller's treatment ([54], pp. 238–241) and in Reichenbach's ([59], p. 176) that the disk observer is *not* employing the space-time metric as his $d\sigma$.

If the angular velocity of the disk is not uniform, then the time-dependence of ω will make the spatial metric $d\sigma$ (and the negative value of the Gaussian curvature) time-dependent as follows: Although any wholly radial spatial interval on K will

remain self-congruent in time amid the variation of ω , the latter's time-dependence will preclude the self-congruence in time of any space interval whose points all have the same radial coordinate ([28], pp. 207 and 209). And the thus time-dependent space metric $d\sigma$ is known to be S -extrinsic to the P_3 of the rotating disk frame from our prior arguments, and extrinsic to the given space-time on the strength of RMH.

Let us now turn to the GTR's dynamical conception of the metric $d\tau^2 = g_{ik} dx^i dx^k$ and, for a given spatial reference frame, of the metric $d\sigma^2 = \gamma_{\alpha\beta} dx^\alpha dx^\beta$. On this conception, the matter-energy distribution is, as it were, Browning's piper who calls the $d\tau^2$ tune via the GTR's field equations for the g_{ik} . We can illustrate this determination of the metrical structure of space-time by reference to the simple case of the continuum of world points on the time-like world-line L of a particle: the matter-energy distribution determines in what particular way an ideal clock attached to that particle, whose increments $d\tau$ assign four-dimensional "lengths" to intervals of L , will *partition* the intervals of L into $d\tau$ -equal equivalence classes. For a given matter-energy distribution, there is thus a unique four-dimensional $d\tau$ -equality and hence a unique partition (or congruence class) for all of the intervals of space-time.

In commenting on Riemann's anticipation of this dynamical conception in section 2(c), (i), I wrote:

As noted plainly by Weyl, one of the things that Riemann is clearly telling us about P -space is this: If "the actual things forming the groundwork of a space . . . constitute a discrete manifold" rather than a continuous one, then there is no corresponding problem at all of going "outside that actuality" to seek the dynamical ground for the CARDINALITY-BASED metric relations among its intervals, metric relations which are "implicit". . . . Unlike continuous P -space, discrete P -space is endowed with (at least) one set of non-trivial metrical relations that are implicit and guaranteed to be independent of contingent dynamical factors (initial or boundary conditions).

Generalizing from Riemann's particular I -intrinsic dyadic property of equicardinality to any S -intrinsic dyadic property, we can say the following: If—contrary to RMH—a continuous four-dimensional PT -manifold (space-time) *did* possess any S -intrinsic rudimentary basis for a nontrivial four-dimensional metric m_{nt} , then Einstein would *not* have needed a contingent dynamical basis for the metrical partitioning of the intervals of the space-time manifold, any more than Riemann needed one for a cardinality-based metric in discrete P -space. As explained by Riemann and elaborated in section 2(c), (i) above, it is precisely because the space-time manifold is devoid of any S -intrinsic rudimentary basis for an m_{nt} metric that external metric standards become necessary, and that the question of a dynamical basis for their metrical behavior then presents itself. If the space-time manifold did possess built-in "residential flats," a metrical description of space-time could be given without concern for the dynamics of the matter distribution.

Hence, following Riemann, I claim that the quest for a dynamical basis for the metric of a continuous P , T , or PT -space is indeed illuminated by the RMH thesis of the intrinsic metric amorphousness of these spaces. And I deny that this illumination is at all lessened by the fact that the matter-energy distribution does determine a *unique* metrical partition of the space-time intervals via the field

equations of the GTR. Nor does it follow from the *uniqueness* of this metrical partition or from the uniqueness of the respective partitions of Newtonian space and time that these partitions are each intrinsic ones. Hence it will not do to point to the uniqueness of the partitions prescribed by either the GTR or Newtonian physics as a basis for inferring their intrinsicity. It is very noteworthy in this connection that Einstein did not regard uniqueness of prescription as proof of intrinsicity. For when speaking of the spatial geometry in a coordinate system associated with a field-producing mass point, he says:

Thus Euclidean geometry does not hold even to a first approximation in the gravitational field, if we wish to take one and the same rod, independently of its place and orientation, as a realization of the same interval ([19], p. 161).

Finally, a comment on Earman's contention that "The GTR implies that there are space-times which have well-defined metrical structures but which are devoid of matter-energy and, hence, of anything that could serve as an extrinsic metric standard" [16]. Earman goes on to conjecture how I would try to reconcile these empty space-times with RMH by saying:

It is not sufficient for Grünbaum to reply that the statements the theory makes about space-time congruence in such cases are to be interpreted by means of counterfactual conditionals, i.e. "If there were extrinsic standards, e.g. atomic clocks and measuring rods present, then the relation . . ."; for since the space-time metric is a dynamical object, the theory implies that if these extrinsic standards were present, the metrical structure would be different, and in order to correct for the perturbation caused by the presence of these extrinsic standards one must know what the structure was in the unperturbed state [16].

I now wish to explain why I maintain that the problem stated here by Earman poses no difficulties for the Riemannian philosophical gloss which I have endeavored to put on the GTR but rather raises philosophical and scientific questions for the GTR itself. And the currently unresolved status of at least the scientific of these questions is acknowledged by practicing relativists. Earman's remarks here call for several comments:

1. Let us suppose with Earman that the so-called empty space-times involved here (i.e. those for which the energy-momentum tensor of the field equations is zero) are to be thought of as *literally* devoid of any of the "test particles" or infinitesimal metrical "test" standards (light clocks, atomic clocks, rods) of which relativity physicists are wont to speak. Then if he wishes to deny me the philosophical invocation of these test bodies as metric standards external to the space-time manifold, it would follow that the GTR *contradicts itself* by containing the following statement: A free particle of non-zero rest mass and also a photon (which has equivalent mass) has a geodesic path in *Minkowskian* space-time. For if cognizance were taken by the theory of the permanent gravitational fields associated with these test bodies, their very presence would have to be held to destroy the Minkowskian character of the space-time by issuing in a nonvanishing four-dimensional Riemann tensor. And in that case, special relativity would cease to qualify as a theory of the behavior of light rays or of any other known physical agencies!

2. Even if these considerations were to be rendered less telling by the pragmatic

appeal to an *approximately* Minkowskian space-time, there is the serious philosophical problem of individuating the elements of an empty space-time which we had occasion to mention several times when discussing the Clifford vision of a space-theory of matter elaborated by Wheeler. (It does not affect the issue here that in the latter vision, space-time is held to be “empty” even when the energy-momentum tensor does *not* vanish.) If there are no *extra*-geochronometric physical entities to specify (individuate) the *homogeneous* elements of space-time and/or of P_3 -space, then whence do these elements of otherwise equivalent punctal constitution derive their individual identities? Must the world points not be individuated before the space-time manifold can even be meaningfully said to have a metric? I see no answer to this question as to the principle of individuation here within the framework of the ontology of the Leibnizian identity of indiscernibles. Nor do I know of any other ontology which provides an intelligible answer to this *particular* problem of individuating avowedly *homogeneous* individuals. And I do not find that Earman came to grips with it in his fetchingly entitled essay “Who is Afraid of Absolute Space?” [17]. In short, if the empty space-times adduced by Earman are held to be *literally* empty in the case of an everywhere vanishing energy-momentum tensor, then I find the event-element ontology of that space-time philosophically unintelligible, quite apart from any need for test bodies to metrize the space-time manifold. I therefore cannot see that an appeal to the metrical structure of a purportedly literally empty space-time can serve to discredit my RMH claim that all of the GTR’s space-times are devoid of *S*-intrinsic m_{nt} metrics.

See Append.
§44 for
amplifying
correction

Correction
in Append.
§44 for the
phrase
“quite
apart from”

3. Let us suppose that no external metric “test” standards *constituted by matter or energy* are being tacitly invoked to metrize the empty space-times adduced here by Earman. It then still does not follow that no kinds of external metric standards are being invoked and that the 4-metric $d\tau$ of the empty space-time solutions of the field equations are *S*-intrinsic. For it may be that in this case, the external metric standard is being supplied by the human imagination via Bridgman’s paper-and-pencil operations.

3. What are the Logical Connections, if any, between Alternative Metrizability, Intrinsic Metric Amorphousness, and the Convention-ladenness of Metrical Comparisons?

(i) *Is there any connection between alternative metrizability and intrinsic metric amorphousness?*

It will be recalled from section 2(c), (iv), 1, that we speak of metrics on a given manifold as “alternative” only if (and if) they differ from one another nontrivially, i.e. other than by a mere scale factor. Thus metrics are alternative only if (and if) they generate different partitions into equivalence classes. And it will likewise be recalled from section 2(c), (i) and (iv), 5, that when we speak of a manifold as being intrinsically metrically amorphous, we always mean that it is devoid of NON-trivial intrinsic metrics, *not* that it lacks trivial intrinsic metrics as well. In section 2(c), (iv), we gave examples of alternative nontrivial *I*-intrinsic *M*-metrics and of alternative *Op*-intrinsic *D_{nt}*-metrics on several arithmetic manifolds. We also exemplified

alternative I -intrinsic M_{nt} -metrics on discrete P -manifolds, and alternative S -intrinsic D_{nt} -metrics on a particular species of these manifolds. Confining our attention entirely to *nontrivial* metrics m_{nt} and D_{nt} when now speaking of alternative metrizable, our proof of the existence of alternative *intrinsic* metrics clearly shows that the alternative metrizable of a manifold certainly is *not* sufficient for its intrinsic metric amorphousness! Indeed, a given manifold S which possesses alternative extrinsic metrics may also be endowed with at least one intrinsic one. And the existence of these alternative extrinsic metrics therefore could hardly serve to establish the intrinsic metric amorphousness of S . (Hereafter, I shall represent the noun phrase "intrinsic metric amorphousness" by "IMA" and the adjectival phrase "intrinsically metrically amorphous" by "*ima*.".) Furthermore, if we assume RMH—as I shall do hereafter unless explicitly stated otherwise—a Riemannian P_3 -space has an *infinitude* of alternative positive definite *extrinsic* metrics $d\sigma^2 = \gamma_{\alpha\beta} dx^\alpha dx^\beta$. Hence, alternative metrizable *as such* cannot serve to distinguish manifolds which are *ima* from those which are endowed with intrinsic metrics. By the same token, as we recall from section 2(c), (iv), 1, and shall note again below, alternative metrizable is *not* sufficient to assure that the ascriptions of equality made by alternative metrics are convention-laden.

What then was the point of my having mentioned alternative metrizable in earlier writings (e.g., [24], Ch. 1), when discussing IMA? There ([24], pp. 44–45), I was addressing myself to a philosophical tradition, typified by the Newtonian position taken by the early Bertrand Russell in his controversy with Poincaré, when Russell, whom I cited there, declared:

It seems to be believed that since measurement [i.e. comparison by means of the congruence standard] is necessary to *discover* equality or inequality, these cannot exist without measurement. Now the proper conclusion is exactly the opposite. Whatever one can discover by means of an operation must exist independently of that operation: America existed before Christopher Columbus, and two quantities of the same kind must *be* equal or unequal before being measured. Any method of measurement [i.e. any congruence definition] is good or bad according as it yields a result which is true or false. Mr. Poincaré, on the other hand, holds that measurement creates equality and inequality. It follows [then] . . . that there is nothing left to measure and that equality and inequality are terms devoid of meaning ([62], pp. 687–688).

Specifically, I was addressing myself to the imperative contained in Russell's statement here, which I analyzed into the following conjunction: (1) If there exist nontrivial intrinsic dyadic properties (of intervals or point pairs), it is the task of metrical comparisons to render them, (2) Newton taught us that there is one and only *one* intrinsic equivalence relation relevant to metrical equality in continuous P_3 -space (if we do not distinguish between intervals and their end-points for this purpose), and hence (3) there is only one *permissible* metrical partition of that space. In the context of the Newtonian imperative here issued by Russell in his debate with Poincaré, the term "alternative metrizable" simultaneously has not only a descriptive but also a *normative* meaning: To say in this context that a space is alternatively metrizable is not merely to make the descriptive statement that alternative metrics can, in fact, be defined on it; instead, it is to say as well *normatively* that the use of these alternative metrics for the purpose of furnishing metrical

comparisons and attributing a metric geometry to the space is *permissible*. I was siding with Riemann against Newton and with Poincaré against Russell by espousing RMH. Hence in that context I was using “alternative metrizable” in the descriptive-cum-normative sense, when I made a point of declaring in opposition to Newton and Russell that continuous P_3 -space is alternatively metrizable. For this was my way of saying to Russell that since RMH is true and the latter P_3 is *ima*, alternative metrics do implement his Newtonian imperative though, of course, only vacuously! And it was in this vein that I stressed ([24], p. 11) that there is no intrinsic basis for singling out one of the infinitude of alternative positive definite Riemannian metrics on P_3 as unique, as Russell demanded of Poincaré.

In the present essay, we have certainly not confined the use of the term “alternatively metrizable” to the restricted descriptive-cum-normative sense. But when we consider below the relative merits of two different metrics on discrete P -space, one of which is extrinsic while the other is intrinsic, we shall allow for an assessment of these relative merits from a normative point of view as well. Indeed, the normative sense of “alternatively metrizable” will also be the one used in our *critique* below of the following thesis: the metrical comparisons made by a theory can be held to be convention-laden *only* if the theory countenances alternative metrics in the normative sense.

It will be noticed that Russell spoke of space intervals, which were at issue in his debate with Poincaré, as either being equal or unequal *simpliciter*. We can now see that even within the confines of the Russellian imperative (1) above, such metrical discourse will become inconsistent IF there are alternative nontrivial intrinsic metrics. Thus, take the two alternative intrinsic metrics μ_1 and μ_2 on one of our species of discrete P -manifold. Since two intervals (or point pairs) which are μ_1 -equal will generally not be μ_2 -equal, an inconsistency will clearly result, if we speak of them as equal or unequal *simpliciter*, while employing *both* of them to render their respective intrinsic equivalence relations. But, of course, no inconsistency ensues from their simultaneous employment for the latter purpose, if we speak of μ_1 -equality and μ_2 -equality respectively. Similarly, take the example of the standard extrinsic metric $ds_1^2 = dx^2 + dy^2$ of the Euclidean plane and then metrize (part of) the same surface alternatively by Poincaré’s non-Euclidean metric $ds_2^2 = (dx^2 + dy^2)/y^2$. We can then speak compatibly of intervals as being ds_1 -congruent but not ds_2 -congruent. And, of course, *if* we use only *one* metric altogether, we can consistently use the term “congruent” or “equal” *simpliciter*, as Russell would have us do.

Russell believed at the time of his controversy with Poincaré that continuous P_3 -space does possess a *unique* intrinsic partition. Hence he urged against Poincaré that we could, in principle, use this unique intrinsic partition to determine whether a transported meter stick (or any other metric standard) “is good or bad according as it yields a result which is true or false.” In other words, when Russell declared that “two quantities of the same kind must *be* equal *or* unequal before being measured,” he apparently appealed to a unique nontrivial intrinsic partition. And he reasoned in the fashion of the Scholium in Newton’s *Principia* as follows: In virtue of the existence of this unique intrinsic partition, it is decidable in principle

whether two intervals are truly spatially congruent (equal) simpliciter or not. Thus, Russell thought that, *given* the intension of "spatially congruent," the properties of Newton's P_3 -space were such that the intension of this predicate determines a unique nontrivial partition as its extension. If the P_3 in question did in fact possess a unique nontrivial intrinsic partition Π , then I would agree that the intension of " x is congruent to y " is indeed such as to determine uniquely the set of ordered pairs of P_3 -intervals corresponding to Π as its extension. But since I have denied the existence of this Π on the basis of RMH, I have maintained the following in earlier writings (e.g. [24], pp. 23 and 27–28, and [24], pp. 248–249): Given the intension which Russell ascribed to the predicate " x is congruent to y ," that intension does *not* suffice to determine a particular partition for any Riemannian P_3 as the extension of that predicate.

This claim prompted a brief criticism by Massey as part of his Panel article ([50], p. 341). But this criticism is superseded by a very illuminating more recent paper of his, devoted specifically to the question "Is 'Congruence' a Peculiar Predicate?" [52]. Among other things, in that paper Massey gives an elaboration which preserves the intent of my intension-extension thesis regarding the congruence predicate. And he likewise shows that the criticisms which Richard Swinburne has made of my views in this connection [65] were addressed to a claim different from my intension-extension thesis. Apart from the matter of Swinburne's reading of my thesis, I trust that the following is already abundantly clear in advance of our impending discussion of the conditions under which ascriptions of metrical equality are convention-laden: The assertion of IMA made by RMH is *not* a semantical truth about the words "length equality" (or " m_{nt} -equality") or " D_{nt} -equality," even if one were to consider as relevant only those m_{nt} and D_{nt} that can be generated by rigid rods and appropriate surveyors' tapes.

Nonetheless, in his review of my *Geometry and Chronometry in Philosophical Perspective* [65], Swinburne says that the only content he can find in the IMA asserted by RMH is the following: It is a *necessary* truth, assured by the meaning of "distance," that metrical determinations in any kind of P -manifold are made by rods and tapes. And hence Swinburne argues that IMA should be ascribed to a *discrete* P -manifold as well, contrary to my Riemannian contention that such a manifold is certainly not *ima*. I shall now cite and then discuss Swinburne's argument in some detail. And I shall do so not only because I hope that my analysis of it may clarify the particular issues raised by him, but for the following additional reason: Swinburne's argument strikes me as a prime recent illustration of how an ordinary language commitment can create the illusion of *a priori* necessity in the case of a merely empirical claim in the manner of Kant's presuppositional method. Swinburne writes:

All that I can understand by the "intrinsic metric amorphousness" of Space is the fact that one measures distances not by counting parts of Space but by applying to Space rods and tapes (and a similar point can be made about Time). This seems a necessary truth. Grünbaum argues that it is a contingent truth which might in fact not hold of Space. He claims that if Space were "granular," it would have an "intrinsic metric." The congruences and metrical attributes of "intervals" would be "intrinsic, being based on the cardinal number of space atoms" (pp. 153f.) But one could only measure distances and establish

congruences by calculating numbers of space atoms on the assumption that such atoms occupied the same volume as each other. But suppose there turned out to be a thousand space atoms between A and B and a million between B and C when, by our ordinary criteria, the distances AB and BC were congruent with each other, what would we say? To my mind, we would regard this as showing that some space atoms take up more room than others. The assumption could thus be shown to be empirically false. So to generalize any "intrinsic" metric possessed by Space or Time would only be such because it satisfied "extrinsic" criteria. The fundamental metric, in other words, would remain the extrinsic one ([65] pp. 310-311).

This set of claims by Swinburne calls for several sets of critical comments.

1. Let us note first that an important *empirical* hypothesis actually underlies the alleged conceptual necessity to metrize any kind of P -space by means of rods and tapes. The pertinent hypothesis was characterized as empirical by Einstein as follows:

All practical geometry is based upon a principle which is accessible to experience, and which we will now try to realise. We will call that which is enclosed between two boundaries, marked upon a practically-rigid body, a tract. We imagine two practically-rigid bodies, each with a tract marked out on it. These two tracts are said to be "equal to one another" if the boundaries of the one tract can be brought to coincide permanently with the boundaries of the other. We now assume that:

If two tracts are found to be equal once and anywhere, they are equal always and everywhere.

Not only the practical geometry of Euclid, but also its nearest generalisation, the practical geometry of Riemann and therewith the general theory of relativity, rest upon this assumption ([20], p. 192).

As I have done previously elsewhere ([28], pp. 272, 277), I shall refer to the empirical assumption just formulated by Einstein as "Riemann's concordance assumption," or, briefly, as "RCA." The latter is, of course, *not* to be confused with RMH.

We see that the empirical truth of RCA plays the following role: It is a necessary condition for the *consistent* use of rigid rods in assigning lengths to space intervals that any collection of two or more initially coinciding unit solid rods of whatever chemical constitution can thereafter be used *interchangeably* everywhere in the P -manifold *independently of their paths of transport, unless* they are subjected to independently designatable perturbing influences. Thus, the assumption is made here that there is a concordance in the coincidence behavior of solid rods such that no inconsistency would result from the subsequent interchangeable use of initially coinciding unit rods, if they remain *unperturbed* or "rigid" in the specified sense. In short, there is concordance among rigid rods such that all rigid unit rods *alike* yield the same metric and the same geometry. It will be recalled that in section 2(c), (i), we had occasion to cite the following comment on RCA by Marzke and Wheeler:

This postulate is not obvious and, in principle, could even be wrong. For example, Weyl once proposed (and later had to give up) a unified theory of electromagnetism and gravitation in which the Riemann postulate was abandoned. In Weyl's theory, two measuring rods, cut to have identical lengths at a point A in space-time, and carried by different routes to a point C , will *differ* in length when they are brought together ([48], p. 58).

There we also noted that Marzke and Wheeler consider "what kind of physics would not be compatible with Riemann's postulate," offer the "Validity of Pauli

Principle as Partial Evidence for Riemann's Postulate" ([48], p. 60), but conclude ([48], p. 61) that "It would be desirable to have a more *decisive* experimental argument for the Riemannian postulate."

In the philosophical literature, Reichenbach has given a vivid description of the kinds of phenomena that would occur if solid bodies were to *violate* RCA, as they do in Weyl's kind of nonRiemannian geometry. This description, which appears in the German original of his classic *Philosophie der Raum-Zeit-Lehre* but not in the 1958 English translation [58], runs as follows:

... whether we could put, say, six chairs in a row into a room would depend on the path by which the chairs were to be brought into the room, and we might perhaps first have to let our chairs make a trip around the world so that the room could accommodate them. By the same token, it would be uncertain whether a visitor could fit onto one of the chairs; this would depend on his prior trajectory. Such states of affairs might perhaps strike us as strange, but they are logically possible; and were they to obtain, humans would surely have come to terms with them ([57], p. 333; translation is mine).

It is, of course, irrelevant to the issue posed by Swinburne that Weyl's nonRiemannian geometry, in which rods are held to violate RCA, did not succeed empirically as a unified theory of electromagnetism and gravitation. Instead, the import of the logical possibility that the hypothesis RCA might be empirically false emerges from the following considerations. Let us assume that in a genuinely discrete *P*-space or in a genuinely granular (atomic, quantized) multiply extended *P*-space there could be physical objects having at least some of the central (essential) properties of the rods or tapes of our actual environment *and* that *one* of these rods could coincide with space intervals under transport in the particular fashion posited by Swinburne. I can imagine states of affairs and physical laws for a granular *P*-manifold in which the metrical findings posited by Swinburne would be physically impossible. Thus, the truth of this assumption cannot be vouchsafed *a priori*.²⁷ Granting this assumption, however, how can Swinburne guarantee as a matter of logical necessity that these rods and tapes in discrete space would exhibit the pattern of lawlike behavior codified by RCA rather than violate it, as the rods of Weyl's geometry are presumed to do? And if the rods in discrete (atomic, granular, quantized) *P*-space did violate RCA, how could they possibly furnish Swinburne with the rod-based distance metric on which he rests his case? How then can Swinburne possibly assure *a priori* that the use of rods (or tapes) would even be feasible to generate a metric in the fashion assumed by him, let alone that they would *have to be* used from conceptual necessity as uniquely legitimate vis-à-vis the alternative, intrinsic cardinality-based metric? And even if rods in a granular manifold did conform to RCA while yielding the measurements posited by Swinburne, why would the cardinality-based intrinsic metric not be at least a *co-legitimate alternative* to the rod-based metric? In other words, why would it have to be the case under the circumstances posited by Swinburne that if the metric generated by rods is indeed extrinsic, then "The fundamental metric . . . would remain the extrinsic one"? In asking this question, I include the if-clause, since an external

²⁷ See the *caveat* in my [28], pp. 152–153 and in my [26], and [27] Ch. III, section 6.

metric standard may, of course, *happen* to generate a metric which is intrinsic, as we recall from section 2(c), (ii) and (iv), 5.

More importantly, whereas circumstances such as the failure of RCA would *eliminate* Swinburne's rods as the metric standard in a discrete *P*-manifold, neither the failure of RCA nor any extrinsic dynamical factors could possibly jeopardize or affect the bases of the *intrinsic* metrics! For the existence of the bases of the non-trivial intrinsic metrics is assured by the very existence of the discrete *P*-manifold itself, as we have emphasized in section 2(c), (i) and (iv), 7.

In short, I maintain that Swinburne has edified the use of the external metric standard constituted by a rod (or tape) into a conceptual necessity for any kind of *P*-manifold by an illicit extrapolation from the following situation: Lacking or at least not knowing of any intrinsic basis for nontrivial metrics in the *P*-manifolds of our environment, we humans have employed rods and taut tapes (though by no means only these!) as external metric standards, because they seem to have the virtues postulated by RCA, and for other reasons as well. And Swinburne's argument also overlooks that even if RCA is always found to hold, refined findings in our actual world may prompt us to give up rods as a metric standard for our *P*-manifolds, much as we gave up the rotating earth as a metric standard for our terrestrial *T*-manifold, as discussed in the **Introduction**.

2. Swinburne asserts "But one could only measure distances and establish congruences by calculating numbers of space atoms on the assumption that such atoms occupied the same volume as each other." And, in the same vein, Earman writes:

Nor is it clear to me what Grünbaum is claiming with his metrical amorphousness thesis. Grünbaum contrasts the (alleged) metrical amorphousness of a continuous space with the intrinsic metrical determinateness of a quantized or granular space by saying that in the latter case there is a built in measure of the interval—the number of granules contained in the interval provides the natural measure of the interval. But this measure rests on the assumption that the granules are all the same size and are all the same distance apart. Perhaps such an assumption can be justified by a physical theory, but so it would seem can assumptions about the metrical structure of continuous spaces [16].

Curiously, A. d'Abro, though apparently an exponent of the RMH implicit in Riemann's Inaugural Dissertation, likewise claims that *special assumptions* must be made to justify the use of the cardinality-based metric on a discrete *P*-space. Thus, d'Abro writes:

... the continuity of space ... precludes us from attributing any absolute meaning to the statement that two lengths situated in different parts of space are equal or unequal. There is no absolute significance in stating that there is as much space between *A* and *B* as between *C* and *D*. We cannot compare lengths by counting the number of points that they contain, since there are just as many points between the extremities of an inch as between the extremities of a mile—an infinite number in either case.

... there is nothing in the continua themselves to suggest any definite metrics or geometry. This is what is meant by saying that space is amorphous and presents us with no means of determining absolute shape and size. ...

This difficulty of counting points might be obviated to a certain degree were space to be considered discrete or atomic; for in that case we might count the atoms of space separating points and thereby establish absolute comparisons between distances. But here again the procedure would be artificial, for it would be nullified unless we were to assume that the

spatial voids separating the successive atoms were always the same in magnitude ([12], p. 40 and 40n.).

I shall address myself to the citations from Swinburne and Earman here, since I do not know why d'Abro speaks of any spatial voids between the elements of an avowedly discrete P -space, unless he is tacitly thinking of that space as embedded in a continuous background space after all. But I cited d'Abro in this connection nevertheless, because he thinks, as Swinburne and Earman do, that there may be something *problematic* about the following in the case of a discrete P -manifold:

(1) the intrinsic cardinality-based M_1 -metric $M(I) = \bar{I}$ ascribes metrical *equality* to all singletons $\{a\}$, (2) the intrinsic cardinality-based D -metric $D_1(a, b) = \bar{I} - 1$, where I is the interval $[a, b]$, ascribes the *same* distance to any *consecutive* pair of granules. And these features of the cardinality-based M -metric and D -metric are claimed to require justification in a way which allegedly renders spurious the distinction between the IMA of continuous P -space and the possession of any intrinsic partition by discrete P -space.

In now explaining why I consider these criticisms as quite unsound, it is grist to my mill rather than a handicap for my position that in section 2(c), (iv), 4, we found certain species of discrete P -manifold to be endowed with nontrivial S -intrinsic metrics *other than* the cardinality-based M_1 and D_1 . Specifically, recall Massey's I -intrinsic "progressive" M -metric M_2 , which is based on the I -intrinsic equivalence relation (dyadic property) of logical identity, and which we did not wish to rule out as trivial even though no two intervals are assigned the same measure by it. In particular, *no* two singletons whatever are assigned the same measure by Massey's M_2 , in striking contrast to M_1 's assignment of the same measure 1 to every singleton. Massey's metric M_2 was relevant to two of our three species of discrete P -manifold. Likewise recall the S -intrinsic metric D_2 on the species having *one* extreme point, which was defined as $D_2(a, b) = |k^2 - l^2|$, where a and b are the k^{th} and l^{th} points respectively. In particular, note that while the cardinality-based D_1 assigns the same distance to *any* consecutive pair of granules, the alternative D_2 certainly does not. The points that I wish to make in rebuttal to Swinburne and Earman can now be made by reference to M_1 and M_2 , but they apply, *mutatis mutandis*, to D_1 and D_2 .

It is clear that the intervals of our species of discrete P -manifold are in fact endowed with the *nontrivial* intrinsic dyadic properties (equivalence relations) which are respectively *rendered by* and constitute the rudimentary bases of M_1 -equality and M_2 -equality. Contrariwise, the W -manifolds to which RMH pertains are *not* similarly endowed. I say that the different I -intrinsic metrics M_1 and M_2 render the *different* I -intrinsic dyadic properties of equi-cardinality and logical identity by means of M_1 -equality and M_2 -equality respectively. In saying this, I do *not* thereby mean to say that the information supplied by one of these metrics may not also be supplied by the other. For we saw in section 2(b), 2, that in addition to being sensitive to differences in identity among intervals, M_2 does specify the cardinality of *every* interval no less than M_1 does, while M_1 is, of course, not sensitive to differences in identity. M_2 has this kind of informational ascendancy over

M_1 , even though M_1 has the distinction of being maximally I -intrinsic in *every* species of discrete P -manifold, while M_2 certainly does not possess the latter virtue in every species of such a manifold on which it is defined. Thus, M_1 and M_2 do not *disagree* in regard to any intrinsic properties of intervals. And hence it would be altogether misleading and wrong-headed to characterize the differences in the numerical assignments which these measure functions make to singletons, for example, as reflecting different “assumptions” about the “sizes” of the singletons in the sense that at least one of them might be mistaken. What *can*, of course, turn out to be false here is the assumption that a given P -manifold is in fact granular to begin with. But *if* P is granular, then it cannot turn out that the measures M_1 , as hypothetically assigned, fail to render the I -intrinsic dyadic property of equi-cardinality of intervals, including of singletons. And this I -intrinsic relational property is *nontrivial* in discrete space while being trivial in dense P -space in our technical sense.

In the case of M_1 and D_1 , Swinburne and Earman speak of the “assumption” that the granules are all the same size and, in the case of consecutive ones, are all the same distance apart. Note that any *nontrivial* M -metric on a *dense* P -manifold likewise assigns the *same* measure $m_{nt}(\{a\}) = 0$ to all singletons $\{a\}$ precisely for the sake of being nontrivial. But nonetheless, no such M -metric m_{nt} on the latter kind of manifold has an *intrinsic* basis, whereas the M -metric M_1 on our discrete space clearly does! (In particular, the m_{nt} generated by Swinburne’s rods and tapes has no intrinsic basis in dense P -space.) This crucial fact is obscured and even belied when Swinburne and Earman speak of M_1 ’s ascription of the measure 1 to all singletons as an “assumption.” For they believe that the use of such a characterization can gainsay the difference between dense and discrete P -manifolds in regard to IMA, which it cannot. Thus, Swinburne’s posited extrinsic metric cannot rob M_1 of its intrinsicity. Nor can it lessen the significance of that intrinsicity. For granular space does have a nontrivial intrinsic story to be told about its intervals (or point pairs) by means of metrical comparisons among them. By contrast, the continuous W -manifolds do not, if RMH is true.

3. Let us return to Swinburne’s question “But suppose there turned out to be a thousand space atoms between A and B and a million between B and C when, by our ordinary criteria, the distances AB and BC were congruent with each other, what would we say?” Under the circumstances posited by Swinburne, two distinct intervals which are equi-cardinal would generally have unequal lengths as measured by rods. Let us grant him the assumptions such as RCA, which we showed to be required by his posit here. And let us assume that the distance metric D_3 generated by rods is indeed extrinsic rather than being the S -intrinsic one $D_2(a, b) = |k^2 - l^2|$ on the unilaterally bounded species of discrete P -manifold. Swinburne had claimed that under the circumstances posited by him, conceptual necessity would require us to use the extrinsic metric D_3 as uniquely legitimate vis-à-vis the cardinality-based M_1 . And I have presented my criticisms of that claim.

But let us now consider the following different, though related contention: The posited metric D_3 (or its M -metrical counterpart M_3) is a *co-legitimate* alternative to the cardinality-based M_1 . What are we to say to this claim of alternative

metrizability in the *normative* sense, given that one of the two alternative metrics is *I*-intrinsic while the other is *S*-extrinsic?²⁸ I shall comment on this question now, and the answer to the associated question as to the conventionality-status of these respective metrics will then be obvious from our next subsection (ii) below.

It seems to me that if a manifold is worthy of metrical characterization, then the object of a metric on the intervals or point pairs of the manifold is to render *intrinsic* facts or tell the *intrinsic* story *in so far as possible*. This norm was contained in the Russellian imperative (1) discussed earlier. Indeed, this desideratum was the basis for our imposition of nontriviality conditions on metrics even in the case of the dense *P*-manifolds, which are devoid of nontrivial intrinsic partitions. With respect to the attainment of this desideratum, there is a clear asymmetry in favor of M_1 between the merits of the intrinsic M_1 and the extrinsic D_3 . The implementation of the normative requirement of therefore using M_1 to the exclusion of D_3 (or of M_3) would, I suspect, also have substantial advantages of descriptive simplicity for the following reason: We can presume that in a world whose physical *P*- (and *T*-) manifolds are indeed granular, the lawful behavior of physical entities would be describable more simply in terms of the intrinsic metric M_1 , rather than being tied in descriptively simple fashion to the D_3 -metric of the transported rods, if there are such at all. And, if the avowed function of any nontrivial metric is indeed to render an existing intrinsic equivalence relation but a particular metric is purported to do so while being extrinsic, then its ascriptions of metrical equality can be indicted as false.

(ii) *What is the connection between the convention-ladenness of metrical comparisons and the extrinsicity of their bases of comparison?*

Suppose the question were put: "Is it a matter of fact or of convention that the platinum-iridium bar which is kept at the Bureau de Poids et Mesures in Sèvres is 1 meter long?" The question is ill-formulated much like "Is Pittsburgh north?", unless it is relativized to a particular domain of fact or states of affairs. Relatively to a certain domain of physical-*cum*-sociological fact or states of affairs, the statement about the meter bar can qualify as a statement of fact, whereas its truth would be held to rest on a convention with respect to any domain of fact that is codified by a physical theory.²⁹ We already had occasion to remark on this relativity of the conventionality-status of some claims to the pertinent set or kind of fact, when we discussed the deliberately far-fetched example of the metrization of the real number continuum on the basis of the *facts* of aesthetic preference furnished by an existing Laplacean demon (cf. section 2(c), (iv), 1).

Now consider some nontrivial metric μ and some statement of metrical comparison *C* which is true with respect to that metric. Given the important differences between various kinds of manifolds in regard to the possession of nontrivial intrinsic metrics, I take it to be very illuminating to make the possession of a basis in INTRINSIC fact the criterion for the factuality, as opposed to convention-ladenness

²⁸ I am indebted to Neal Grossman and Clark Glymour for having put a question to this effect to me in 1968.

²⁹ For a further discussion of this statement, see my [28], pp. 270–281.

of C , where the sentence " C has a basis in intrinsic fact" is to be understood as follows: Either the truth of C follows from the nontriviality conditions for the metric or its truth is vouchsafed by a component of the metric which has an intrinsic basis. Thus, I take the existence of an intrinsic basis for the equality-component of a nontrivial metric μ as the kind of fact pertinent to and decisive for the factuality or convention-ladenness of the ascriptions of metrical equality made by μ . And similarly for the ordinal and ratio components of the metric.

I stressed in the overall **Introduction** to this essay, and Reichenbach, among others, emphasized repeatedly earlier, that there are indeed facts of coincidence and concordance with respect to which the truth or falsity of a claim about the length of a body in meters (to within a certain precision) is most certainly *not* a matter of stipulation. But such facts of coincidence and concordance involve entities *external* to our P -manifolds. Hence the latter facts are not comprised in the domain of intrinsic fact or states of affairs which we have taken to be the one pertinent to the ontological status of metrical comparisons.

Indeed, it is with respect to the domain of **INTRINSIC** facts or states of affairs that equivalence obtains among alternative descriptions that Reichenbach calls "equivalent" in his theory of equivalent descriptions as applied to metric descriptions. For example, consider *any* two different nontrivial extrinsic metrics ds_1 and ds_2 defined on the intervals of a continuous P_3 -manifold. For concreteness, let ds_1 and ds_2 be alternative positive definite Riemannian metrics on P_3 . For *each* of these two metrics, there will, of course, be a function from the measures of the intervals to their cardinality: If the measure ds of an interval is zero, then it is a singleton, and if ds is positive, then the interval has the cardinality of the continuum. Hence *both* ds_1 and ds_2 will render the intrinsic facts of equicardinality: In the case of either metric, any two intervals which each receive the measure zero will both be singletons, and any two intervals which each receive positive measures (be these measures different or the same in value) will be equi-potent. Hence the extrinsic metrics ds_1 and ds_2 equivalently describe the intrinsic facts of equicardinality of intervals. The two nontrivial metrics generate *different* partitions. In *either* partition, all members of the same cell of the partition will have the intrinsic dyadic property of equi-cardinality. But this latter fact about the members of each *cell* does not, of course, furnish an intrinsic basis for the partition into cells, since equicardinality provides no basis at all for placing nondegenerate intervals into *different* cells of the partition. The set of degenerate intervals will be a member of every nontrivial partition.

Suppose that one postulates, as in Milne's cosmology, that atomic clocks and astronomical ones effect different nontrivial partitions in the time continuum, their respective time coordinates being logarithmically related. Then one can say with Reichenbach that the metrical description in terms of atomic time t and the alternative metrical description in terms of the astronomical time τ are logically equivalent with respect to "the facts," *provided* that the latter term refers to *intrinsic* facts. For with respect to other kinds or domains of fact, metric descriptions that Reichenbach calls "equivalent" can surely be logically *inequivalent*. Milne's t -time description does not, *taken by itself*, give the same *extrinsic* information as his τ -

time description, taken by itself, although the addition of the statement of the exact logarithmic relation between the two time scales makes it possible to *correlate* the two descriptions, or to "translate" one into the other in that sense. Thus, this intertranslatability is tantamount to logical equivalence of the two descriptions with respect to all matters of intrinsic fact, not with respect to all facts simpliciter. By contrast, note that alternative nontrivial *intrinsic* metrics on a discrete P -manifold, whose comparisons render different intrinsic dyadic properties, obviously are not equivalent with respect to intrinsic facts.

I should explain why I speak of a metrical comparison C generated by a nontrivial extrinsic metric as "convention-laden" rather than as outright "conventional." I have also used the locution "conventional ingredient", as when I said in the **Introduction** that ascriptions of metrical equality and inequality generated by a nontrivial extrinsic metric contain an important conventional ingredient, instead of saying that these ascriptions are simply conventional. The reason is that by satisfying the necessary conditions for nontriviality, the metrical comparisons generated by a nontrivial extrinsic metric are sensitive to certain intrinsic facts. For example, the fulfillment of the first of our necessary conditions for nontriviality (cf. section 2(b), 2) by an extrinsic metric assures its sensitivity to the intrinsic dyadic property of proper inclusion when possessed by two closed intervals. And we saw that a nontrivial extrinsic M -metric on continuous P -space does convey information on the intrinsic equicardinality of intervals. Thus, the comparisons generated by a nontrivial extrinsic metric do convey intrinsic information.

I should also point out how one of my other formulations in the **Introduction** and in earlier writings is to be construed in terms of the apparatus of articulations presented in the present essay. I said there that if, say, a transported meter stick (or other external entity) is chosen as the metric *standard* to metrize a continuous P_3 -manifold, then the assertion of the "self-congruence" of the meter stick (and of its respective calibrated parts) under transport is conventional. This statement of mine is to be construed as an ellipsis for saying that since the nontrivial partition generated by meter-equality is extrinsic, ascriptions of meter-equality and inequality are convention-laden. This statement requires a very minor qualification or exception in regard to the M -metrical equality of singletons, the D -metrical equality of the *special* ordered point pairs (a, a) and (b, b) and other metrical ascriptions which follow from the nontriviality conditions. To make the first of these exceptions, which is guaranteed by the nontriviality conditions, quite intuitive, note the following:

If the measure function generated by a meter stick is defined *only* on *intervals*, as in the case of a length function, then zero will be assigned to singletons and *only* to singletons. Hence when a length function based on the meter stick ascribes metrical equality to two intervals in the particular case in which both intervals are of *zero* length, that special kind of comparison C is true if and *only if* the two intervals are equicardinal. Thus, assertions of meter length equality *zero*, and equi-distance *zero*, do have an intrinsic basis. And consequently this unexciting class of assertions is one of the exceptional ones by not being convention-laden. Similar qualificatory remarks apply to a meter stick metric defined on *only* the

intervals of a dense denumerable P -space. Having said this, I shall spare myself the repetition of this qualification in subsequent remarks, although it is to be understood as being asserted tacitly. But recall that for the *indefinite* R -metrics on relativistic space-times, it is *not* the case that two intervals both receive the value zero *only* if they are both singletons (section 2(b), 5). Hence no corresponding qualification is needed for these indefinite R -metrics.

I have been speaking of the conditions under which metrical *comparisons* are convention-laden rather than of the metric itself as being convention-laden. To have done the latter without greater specificity would have been undesirable. For the different components of a given metric may differ in regard to their possession of an intrinsic basis. In section 2(c), (ii) and (iv), 4, we encountered metrics whose equality component does have an intrinsic basis but whose ratio-component, for example, does not. In such a case an ascription of equality made by the metric would *not* be convention-laden, while any comparison asserting a metrical ratio different from 1 would be convention-laden. It is clearly desirable to allow for the latter discrimination, which would not be achieved by merely characterizing the metric itself as convention-laden. If a given kind of metrical comparison is *not* convention-laden, I speak of the comparison in question as being “nonconventional.” For reasons explained in section 2(c), (iv), 4, I leave the possible unit-intrinsicality of a metric out of account and speak of a metric as “maximally” intrinsic (for any given species of intrinsicality), iff it is ratio-intrinsic, i.e. iff all of its components apart from the unit component have an intrinsic basis. Correspondingly, I refer to a maximally intrinsic metric as maximally nonconventional, even though two *trivially different* cardinality-based metrics on discrete P -space may differ in regard to unit-intrinsicality. But if one did wish to take cognizance of unit-intrinsicality when present, one could use the term “entirely nonconventional” in a special technical sense to characterize a metric which is both ratio-intrinsic and unit-intrinsic, as distinct from one which is only ratio-intrinsic.

It follows from RMH that for any nontrivial metric μ_{nt} on a W -manifold, all of the components of μ_{nt} are devoid of an intrinsic basis. Therefore, any μ_{nt} -based metrical comparisons pertaining to a W -manifold are convention-laden, with the unexciting exceptions that follow from the nontriviality conditions, such as the case of equality in regard to *vanishing* size with respect to certain kinds of metrics. And the convention-ladenness of these μ_{nt} -based metrical comparisons is quite independent of whether μ_{nt} is generated by only one kind of external standard (e.g. rigid rods) or by a whole cluster of kinds of standards (rods, radar ranging, etc.). Thus, the claim I made to this effect in the **Introduction** is vindicated here. On the other hand, the nontrivial cardinality-based metric M_1 on discrete P -space is maximally nonconventional, since all of the components of M_1 , with the possible exception of its unit-component, have intrinsic bases. And hence all M_1 -generated metrical comparisons of intervals of discrete P -space are nonconventional, as are those of the cardinality-based D_1 on that space. By contrast, the metrical comparisons of Swinburne’s extrinsic D_3 are convention-laden.

We saw in subsection (i) that the alternative metrizable of a manifold S is not sufficient for S ’s being *ima*. For S may even be endowed with alternative *intrinsic*

metrics. And it is evident now that quite generally the existence of alternative metrics on S likewise is *not* sufficient to establish the convention-ladenness of any metrical comparison generated by one or another of the alternative metrics. For convention-ladenness arises from the lack of an intrinsic basis and *not* from the existence of an alternative metric! In this connection, the reader is asked to recall from section 2(b), 2, and section 2(c), (iv), 1, our account of the complete compatibility of this claim with the fact that we may have a choice among alternative metrics.

These considerations enable us to evaluate the characterization of conventionality given by Earman on which he rests his critique of important facets of my position. He says:

the question of the conventionality of spatial and temporal congruence . . . amounts to the question of whether within a given frame of reference one can with equal factual legitimacy adopt alternative measures of spatial and temporal intervals which yield different congruence classes [16].

Although Earman is using this formulation as a basis for a discussion of my views in the particular context of the GTR, it merits comment from a less restrictive perspective. And I shall include remarks specifically addressed to the GTR context at suitable places below. The crucial phrase in Earman's statement is "one can with equal factual legitimacy." The terms "can" and "legitimacy" may have both descriptive and normative meanings, and the normative use of these terms may pertain to what a theory countenances normatively as a basis for the *particular formulations* by means of which it may codify the facts asserted by it. In Section 4 of his paper, Earman does indeed invoke the normative meaning of these terms, and I shall defer consideration of the normative issue until sub-part (iii) of the present Section 3.

Let us, therefore, now construe Earman's statement as confined to the descriptive sense of "can with equal factual legitimacy." Then we can show that he is *not* addressing himself to my conception of convention-ladenness, when he treats the existence of factually co-legitimate alternative metrizations of the same P -manifold as *sufficient* for the convention-ladenness of the equality component of each of the alternative metrics. For let our spatial reference frame be the unilaterally bounded discrete P -manifold of our section 2(c), (iv), 4(β). Then the cardinality-based D -metric D_1 , and the alternative D -metric $D_2(a, b) = |k^2 - l^2|$ discussed there effect different partitions (congruence classes). But they are both S -intrinsic and just render different intrinsic equivalence relations. Clearly, D_1 and D_2 are *factually co-legitimate*! Both being intrinsic, neither D_1 -equality nor D_2 -equality is convention-laden. By Earman's cited criterion, the equality component of each of these metrics should be convention-laden.

To say that Earman is not employing my criterion for convention-ladenness here is, of course, *not* to say that there are no interesting cases in which alternative metrics on a given manifold do each have a convention-laden equality component: In the time continuum of Milne's cosmology, the two different nontrivial partitions associated with his aforementioned two time-scales are each extrinsic. And thus

the metrical equalities generated by these two time scales are each convention-laden.

It will be remembered that when I discussed the meaning of RMH and the evidence supporting it in section 2(c), (i) and (iv), 5, I was careful *not* to assume that *only* intrinsic properties can be topological. Hence suppose that ascriptions of topological properties can be divided into two classes such that those in the first class do have intrinsic bases while those in the second do not. Then I would certainly be prepared to characterize them in regard to convention-ladenness along lines similar to those in the account I gave for metrical comparisons. And I trust it is clear that in either the metrical or topological case, convention-ladenness and nonconventionality of an ascription pertains, at least in the first instance, to the *ontological* status of the ascription and not to its epistemic status. Just as I argued in the present essay in the Reichenbachian tradition that the nonconventionality of a metrical comparison of intervals or point pairs of a manifold will depend on the facts as to the intrinsic structure of the manifold, so also I have recently argued elsewhere in concert with several colleagues that an analogous observation applies to ascriptions of simultaneity [33]: If the pertinent facts are of the kind assumed in Newton's world, then the ascription "noncoincident events x and y occupy the same temporal locus in a common system of temporal order" is nonconventional, but if the pertinent facts are of the different sort postulated by Einstein, then this ascription of simultaneity is convention-laden.

As far back as during the Middle Ages, Walter Burley wrote: "... in a continuum [which is a P -manifold] there is no primary and unique measure according to Nature, but only according to the institution of men" ([11], p. 103). I take this to be an early adumbration of part of the substance of RMH, and of the account of the convention-ladenness of metrical comparisons which I have presented along with RMH.

(iii) *Is normative alternative metrizable either sufficient or necessary for the convention-ladenness of metrical comparisons?*

Consider a physical theory dealing with the unilaterally bounded discrete P -manifold of our section 2(c), (iv), 4(β). We saw in the preceding sub-part (ii) that both of the two D -metrics D_1 and D_2 on that manifold discussed there yield *nonconventional* ascriptions of metrical equality. Now suppose that a certain physical theory pertaining to that P -manifold were to countenance the use of both D_1 and D_2 normatively so as to render each of the intrinsic dyadic properties (equivalence relations) which constitute the respective rudimentary intrinsic bases of these two metrics. We can then say that there is *normative* alternative metrizable in the given physical theory in the following sense: The theory in question countenances the use of both of two alternative metrics, each metric having a purpose of its own to render a certain intrinsic equivalence relation, for example. It is then evident from the nonconventionality of both D_1 -equality and D_2 -equality that normative alternative metrizable is *not* a sufficient condition for the convention-ladenness of the equality-component of at least one of the alternative metrics.

Earman has claimed by special reference to the GTR that normative alternative metrizability is a *necessary* condition for the convention-ladenness of metrical comparisons. And he has adduced this premiss to do both of the following: (1) To object to the conventionality-status which I assign to the time-dependent species of spatial metric $d\sigma$ discussed in section 2(c), (iv) 7, and (2), to contend that my own avowal of the *normative* uniqueness of the GTR's metrical partition of space-time refutes my RMH-based claim that the GTR's ascriptions of space-time equality are convention-laden. I shall now discuss these two arguments of Earman's in turn.

1. Speaking of my example of a time-dependent spatial metric $d\sigma$ on the rotating disk, Earman writes:

The congruence classes of infinitesimal spatial intervals will be different at different times, but these different classes do not represent alternative congruences in any sense which supports the conventionality thesis since at no instant is there any freedom in the choice of the measure of these intervals in the chosen frame. If a proponent of the thesis of the conventionality of temperature should claim that it is a matter of convention to assign a temperature to a given region of the earth, and all he meant by this claim was that the temperature of the region is at any given moment empirically well defined but is different at different moments, we would not be overly eager to hear more about his thesis [16].

It seems to me that Earman's temperature analogy is infelicitous. For I proved in section 2(c), (iv) 7, that any time-dependent metric on a *spatial* P_3 -manifold *must* be *S*-extrinsic. And it then follows from our preceding sub-part (ii) that the equality-component of any such metric is convention-laden, if the assignments of equal and unequal lengths to spatial intervals at different times are to rest on a common basis of comparison. Earman nonetheless thinks that the different, time-dependent partitions "do not represent alternative congruences in any sense which supports the conventionality thesis since at no instant is there any freedom in the choice of the measure of these intervals in the chosen frame." Here he is invoking his contention that *normative* alternative metrizability is a necessary condition for the convention-ladenness of metrical comparisons. But our analysis has shown that this claim is false: It is not true that the metrical comparisons generated by a given metric must be nonconventional, just because the metric in question is the only one normatively countenanced by the theory in which it is used. For we saw that convention-ladenness arises from the lack of an intrinsic basis and hence cannot require normative alternative metrizability. Earman's claim that normative alternative metrizability is a *necessary* condition for the convention-ladenness of metrical comparisons is just as unsound as the assertion that it is a sufficient condition, an assertion whose falsity we established at the outset of the present sub-part (iii).

As for his temperature analogy, its inappositeness seems to me to arise from the following facts. As I showed in section 2(c), (iv) 7, all of the *intrinsic* properties of a spatial manifold P_3 and of any of its intervals obtain *timelessly*, and extrinsicity is interestingly relevant to convention-ladenness in the case of such a manifold. By contrast, no corresponding statement of timelessness holds for the intrinsic properties of a body which are relevant to its temperature, if there are such for a given concept of temperature. And it is unclear that the intrinsic properties of a body can

generally play as interesting a role as the intrinsic properties of a *P*-manifold, *T*-manifold or *PT*-manifold. Yet I can think of cases in which the *intrinsic* properties of a body might play such a role, but *in contrast to those of P-space*, are lawfully *time-dependent*. For example, consider the number of radioactive atoms of a given specimen of radioactive substance which are undecayed at a given time, or rather the ratio of the number of decayed atoms to the number of undecayed ones in the specimen at a given time. This ratio could interestingly be thought of as an *intrinsic* property of the given specimen, and yet its lawful time-dependence is specified by the Rutherford-Soddy law.

2. Earman again invokes the false assertion that normative alternative metrizability is a necessary condition for convention-ladenness of metrical comparisons, when he makes the following allegation: My own avowal of the *normative* uniqueness of the GTR's metrical partition of space-time (for a given matter-energy distribution) refutes my RMH-based claim that the GTR's ascriptions of space-time equality are convention-laden. He presents the gravamen of this charge by arguing as follows:

. . . metrical amorphousness is said to theoretically allow and to legitimate a conventional choice of the congruence relations for disjoint intervals of a continuous physical space or time.

It follows that Newton's [*philosophical!!*] theory of space and time could not possibly have been right; and Grünbaum himself presses this point. . . . But if Newton could not possibly have been right, then neither could Einstein, for as Putnam³⁰ points out, if Grünbaum's argument works at all for space and time, it should also work for the space-time continuum. But as Grünbaum himself acknowledges, the GTR does not leave any room for conventionality with respect to the congruence of space-time intervals; nor does it leave any room for conventionality with respect to the congruence of spatial and temporal intervals, or so I have argued. [16].

I certainly did avow that the GTR effects a normatively unique partition of its space-time intervals for a given matter-energy distribution. But I have *not* thereby acknowledged that "the GTR does not leave any room [in my descriptive sense] for conventionality with respect to congruence of space-time intervals": To maintain that a certain normative requirement of a theory is convention-laden is surely *not* to say that the theory *cannot* adopt that requirement to the permanent exclusion of all alternatives to the object, i.e. the 4-metric, governed by the requirement! Suppose, *per impossibile*, that it had been a normative requirement in all human societies ever since the rise of *homo erectus* half a million years ago that the meter and its decimal subdivisions be the only units employed in all practical affairs and in any calculations made in any physical theory ever devised by man. Could this impugn the contention that the ascription of the length 1 meter to the platinum-iridium bar in Sèvres is convention-laden as a statement of our present-day physical theory?

See Append.
§45 for an
amendment

4. Intrinsicness and Extrinsicness of a Relation on a Manifold

Section 2(c) presented an articulation of the distinction between intrinsic and extrinsic metrics, which I had only adumbrated in earlier writings. In a paragraph of *Philosophical Problems of Space and Time* ([24], pp. 214–215), I only sketched the distinction between the intrinsicity and extrinsicity of a relational property

³⁰ At this point, Earman cites Putnam's [56].

on a manifold. And I did so by reference to the following particular example: The manifold was the complete topological Euclidean straight line, and the kinds of relational properties of *elements* of that manifold whose intrinsicality-status I considered there were Huntington's betweenness (cf. section 1(a)) and a serial relation. The present section 4 is designed to replace my very inadequate earlier illustrative sketch by general and more precise formulations, and to explain the results which were anticipated in section 1(a).

In *Philosophical Problems of Space and Time* [24], I was concerned with the conditions under which a serial relation (relational property) on a manifold is or is not intrinsic, because I was interested in relating that distinction to the difference between two kinds of time continua: a *T*-space which is endowed with an "arrow" and one which is not. And in section 1(a) of the present essay, I alluded to irreversible processes when speaking of the intrinsicality of a serial relation on a continuous *T*-space which is not unilaterally bounded. As indicated by that allusion, I believe that the more precise formulation of the intrinsicality of a serial relation to be given here can be made illuminatingly relevant to the discussion of time's arrow. But I shall defer an account of that temporal application to the second installment (**Part B**) of this essay, where I shall reply to Earman's critique ([14], pp. 543–549 and [15], pp. 273–295) of my writings on the anisotropy of time. For the statement of the particular *application* of the material in the present section 4 to the time continuum is not at all essential to the presentation of our definitions. And I have deferred to **Part B** all replies to critiques, unless such replies were directly helpful in giving the new exposition to which the present **Part A** is devoted.

I shall eschew redeveloping here notions characterized in the discussion of intrinsic metrics. Thus, for example, the reader is asked to recall from section 2(c), (ii) the remarks there about my intensional construal of what constitutes a property or relation, and the comments about the notion of a general property and about logical dependence as bases for the following definitions.

(1) *Given* the elements of a manifold *S*, a monadic property *P* of the elements of *S* is intrinsic to them, iff *P* is a *general* property whose possession by an element depends on the existence of no entities other than the elements of *S*.

(2) Similarly for an *n*-adic intrinsic property *P_n* of the elements of *S*.

(3) An *n*-adic property *P_n* of the elements of *S* is extrinsic to them, iff it is not intrinsic to them.

For example, the triadic property *B* of betweenness of our section 1(a), possessed by the elements of singly extended *P*-manifolds which are discrete, dense denumerable or linear continuous, is seen to be an *intrinsic* ternary property by definition (2). Hence we can also see that for any such *P*-manifold which is *unilaterally bounded*, BOTH of the following two properties of its elements are *intrinsic*: Firstly, the monadic property of being an extreme element defined by

$$x \text{ is an extreme element} =_{\text{Def.}} \sim (\exists y)(\exists z)B(yxz),$$

(where *B* is the aforementioned intrinsic triadic property of betweenness), and secondly, the *serial* dyadic property *P* ("precedes") defined by

$$P(x, y) =_{\text{Def.}} (\exists z)[(z \text{ is an extreme element}) \cdot \{(x = z) \cdot (y \neq z) \vee B(zxy)\}]$$

(4) An n -adic intrinsic property P_n of the elements of a manifold S is said to be intrinsic to S , and correspondingly for an extrinsic property.

We found in section 2(c), (ii) that it may be possible to specify an intrinsic metric by means of external entities or devices, so that the *description* of the metric will be extrinsic even though the metric is actually intrinsic. In somewhat analogous fashion, if an n -adic property qualifies as intrinsic to a manifold S , the *extension* of that property might also be specifiable by means of devices external to S . For example, suppose that a unilaterally bounded discrete manifold, such as a string of beads having exactly one extreme bead, were viewed by an outside onlooker for whom the extreme bead-element is the leftmost one. Then, although the serial property $P(x, y)$ defined above is intrinsic to this kind of manifold, the *extension* of that dyadic property could also be specified by means of the intension of the term "to the left of" as understood from the particular perspective of that external viewer. But this possibility does not, of course, militate against the intrinsicity of the stated serial property $P(x, y)$, which is *intensionally* different from the property of the viewer's "being to the left of." But if no intrinsic relational property is known of which a given relation-in-extension on a manifold is indeed the extension, this fact creates the presumption that the manifold may well be devoid of such an intrinsic property.

As we noted in section 1(a), definitions of serial order properties for all manifolds which are systems of betweenness (in the sense of Huntington {[39], pp. 2-3} or in the sense of Tarski {[67], p. 17}) have been given by Huntington, Tarski, and by Massey in his Panel article ([50], p. 339). But, as we remarked there in section 1(a), the constitution of the serial order properties specified by these definitions does involve particular individuals in an essential way. And since these properties are therefore not general, they fail to qualify as intrinsic to manifolds which are systems of betweenness.

A P -space having *exactly one* extreme element is, of course, only a very special kind of species of P -model of the postulates for betweenness. And we noted above that this particular kind of P -space is endowed with an intrinsic serial property $P(x, y)$ whose definition makes essential use of the possession of an extreme element by that P -space. What of the case of a singly extended discrete, dense denumerable, or linear continuous P -space K having exactly *two* extreme points? Consider a binary predicate " Pxy " definable so that the predicate " B " denoting the intrinsic triadic property B is the only nonlogical constant in the definiens. Then Massey has shown that for any space of the type K , there is no " Pxy " of this kind such that " Pxy " determines a serial ordering of the points of K .

More generally, it can be presumed that only the unilaterally bounded species of singly extended discrete, dense denumerable or linear continuous P -space is endowed with an intrinsic serial property.

By contrast, consider the two *arithmetic* manifolds of the natural numbers and of the real numbers in their familiar customary construal. It will have been noted that in definitions (1) and (2) of the present section 4, no less than in definitions (1) and (12) for metrics, I took the elements of the manifolds as *given* entities. Hence, when I now discuss serial properties which are intrinsic to these two manifolds,

certain things should be borne in mind: I am not pretending to offer any philosophy of arithmetic but am using arithmetic manifolds only illustratively, and it suffices for my purposes if the kinds of numbers I am about to use in my illustrations are obtained by what Russell has called the "method of theft" rather than by "honest toil." But my procedure certainly does not exclude the method of honest toil!

In the domain of *natural* numbers, the following serial property "less than" qualifies as *intrinsic* by our definitions:

$$x < y =_{\text{Def.}} (\exists z)[\sim(w)(w + z = x) \cdot \{x + z = y\}]$$

And in the domain of *real* numbers (negative, zero and positive), the "less than" property specified by

$$X < Y =_{\text{Def.}} (X \neq Y) \cdot (\exists Z)(X + Z^2 = Y)$$

is an intrinsic one. We omit the more complicated case of "less than" for the rationals.

Note the contrast in regard to being endowed with an intrinsic serial order property between the linear continuous *P*-manifold of the complete Euclidean spatial line and the linear Cantorean continuum of the real numbers (negative, zero and positive). This contrast points up anew the important role played by the constitution of the elements of a manifold in our "material" (non-formal) conception of intrinsicity.

REFERENCES

- [1] Adler, R., Bazin, M., and Schiffer, M., *Introduction to General Relativity*, McGraw-Hill, New York, 1965.
- [2] Anderson, J. L., *Principles of Relativity Physics*, Academic Press, New York, 1967.
- [3] Arzelès, H., *Relativistic Kinematics*, Pergamon Press, New York, 1966.
- [4] Bergmann, P. G., "The General Theory of Relativity," in *Handbuch der Physik*, vol. IV (ed. S. Flügge), Springer, Berlin, 1962, pp. 203-272.
- [5] Bondi, H., *Cosmology*, 2nd ed., Cambridge University Press, Cambridge, 1961.
- [6] Bourbaki, N., *Éléments d'Histoire des Mathématiques*, Hermann, Paris, 1969.
- [7] Bridgman, P. W., "Some implications of Recent Points of View in Physics," *Revue Internationale de Philosophie*, vol. III, No. 10 (1949).
- [8] Cech, E., *Point Sets*, Academic Press, New York, 1969.
- [9] Clifford, W. K., *The Common Sense of the Exact Sciences*, Dover Publications, New York, 1955.
- [10] Clifford, W. K., "On the Space Theory of Matter," in *The World of Mathematics*, vol. 1 (ed. J. R. Newman), Simon and Schuster, New York, 1956, pp. 568-569.
- [11] Crombie, A. C., *Robert Grosseteste and the Origins of Experimental Science*, Oxford University Press, Oxford, 1953.
- [12] d'Abro, A., *The Evolution of Scientific Thought from Newton to Einstein*, Dover Publications, New York, 1950.
- [13] Demopoulos, W., "On the Relation of Topological to Metrical Structure," in *Minnesota Studies in the Philosophy of Science*, vol. 4 (eds. M. Radner and S. Winokur), 1971.
- [14] Earman, J., "Irreversibility and Temporal Asymmetry," *Journal of Philosophy*, vol. 64 (1967), pp. 543-549.
- [15] Earman, J., "The Anisotropy of Time," *Australasian Journal of Philosophy*, vol. 47 (1969), pp. 273-295.
- [16] Earman, J., "Are Spatial and Temporal Congruence Conventional?" (forthcoming).
- [17] Earman, J., "Who is Afraid of Absolute Space?", *Australasian Journal of Philosophy*, December, 1970.

- [18] Eddington, A. S., *The Mathematical Theory of Relativity*, Cambridge University Press, Cambridge, Mass. 1952.
- [19] Einstein, A., "The Foundations of the General Theory of Relativity," in *The Principle of Relativity, A Collection of Original Memoirs*, Dover Publications, New York, 1952.
- [20] Einstein, A., "Geometry and Experience," in *Readings in the Philosophy of Science* (eds. H. Feigl and M. Brodbeck), Appleton-Century-Crofts, New York, 1953, pp. 189–194.
- [21] Eisenhart, L. P., *Riemannian Geometry*, Princeton University Press, Princeton, 1949.
- [22] Feinberg, G., "Pulsar Test of a Variation of the Speed of Light with Frequency," *Science*, vol. 166 (1969), pp. 879–881.
- [23] Grünbaum, A., "Geometry, Chronometry, and Empiricism," in *Minnesota Studies in the Philosophy of Science*, vol. III (eds. H. Feigl and G. Maxwell), University of Minnesota Press, Minneapolis, 1962, pp. 405–526.
- *[24] Grünbaum, A., *Philosophical Problems of Space and Time*, Alfred Knopf, New York, 1963, and Routledge and Kegan Paul Ltd., London, 1964.
- [25] Grünbaum, A., "The Falsifiability of a Component of a Theoretical System," in *Mind, Matter and Method: Essays in Philosophy and Science in Honor of Herbert Feigl* (eds. P. K. Feyerabend and G. Maxwell), University of Minnesota Press, Minneapolis, 1966, pp. 273–305.
- [26] Grünbaum, A., *Modern Science and Zeno's Paradoxes*, Wesleyan University Press, Middletown, Conn., 1967. A revised edition is listed next.
- [27] Grünbaum, A., *Modern Science and Zeno's Paradoxes*, Allen and Unwin Ltd., London, 1968.
- [28] Grünbaum, A., *Geometry and Chronometry in Philosophical Perspective*, University of Minnesota Press, Minneapolis, 1968.
- [29] Grünbaum, A., "Are Physical Events Themselves Transiently Past, Present and Future? A Reply to H. A. C. Dobbs," *The British Journal for the Philosophy of Science*, vol. 20 (1969), pp. 145–162.
- **[30] Grünbaum, A., "Can We Ascertain the Falsity of a Scientific Hypothesis?" *Studium Generale*, vol. 22 (1969), pp. 1061–1093. This essay will also appear (revised) in *Observation and Theory in Science*, Johns Hopkins Press, 1971.
- [31] Grünbaum, A., "Reply to Hilary Putnam's 'An Examination of Grünbaum's Philosophy of Geometry'," in *Boston Studies in the Philosophy of Science*, vol. V (eds. R. S. Cohen and M. W. Wartofsky), Reidel Publishing, Holland, 1969, pp. 1–150.
- *[32] Grünbaum, A., "Simultaneity by Slow Clock Transport in the Special Theory of Relativity," *Philosophy of Science*, vol. 36 (1969), pp. 5–43.
- [33] Grünbaum, A., Salmon, W. C., van Fraassen, B. C., and Janis, A. I., "A Panel Discussion of Simultaneity by Slow Clock Transport in the Special and General Theories of Relativity," *Philosophy of Science*, vol. 36 (Mar., 1969), pp. 1–81.
- [34] Halmos, P. R., *Measure Theory*, Van Nostrand, New York, 1950.
- [35] Hasse, H. and Scholz, H., *Die Grundlagenkrisis der griechischen Mathematik*, Pan-Verlag, Charlottenburg, 1928.
- [36] Hempel, C. G., "On the 'Standard Conception' of Scientific Theories," in the *Eisenberg Memorial Lecture Series, 1965–66* (ed. R. Suter), Michigan State University Press, East Lansing, 1970.
- [37] Heyting, A., *Axiomatic Projective Geometry*, North-Holland Publishing, Amsterdam, 1963.
- [38] Hobson, E. W., *The Theory of Functions of a Real Variable*, vol. I, Dover Publications, New York, 1957.
- [39] Huntington, E. V., "Inter-Relations Among the Four Principal Types of Order," in *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 1–9.
- [40] Huntington, E. V., *The Continuum and Other Types of Serial Order*, 2nd ed., Harvard University Press, Cambridge, 1942.
- [41] Hurewicz, W. and Wallman, H., *Dimension Theory*, Princeton University Press, Princeton, 1941.
- [42] James and James, *Mathematics Dictionary*, 3rd ed., Van Nostrand, Princeton, 1968.
- [43] Janis, A. I., "Synchronism by Slow Transport of Clocks in Noninertial Frames of Reference," *Philosophy of Science*, vol. 36 (Mar., 1969), pp. 74–81.
- [44] Landau, L. and Lifschitz, E., *The Classical Theory of Fields*, Addison-Wesley Press, Cambridge, Mass., 1951.

- [45] Leonard H. S. and Goodman, N., "The Calculus of Individuals and Its Uses," *Journal of Symbolic Logic*, vol. 5 (1940), pp. 45–55.
- [46] Lipschutz, S., *General Topology*, Schaum Publishing, New York, 1965.
- [47] Luria, S., "Die Infinitesimaltheorie der antiken Atomisten," *Quellen und Studien zur Geschichte der Mathematik, Astronomie, und Physik*, Abteilung B, Studien II, Berlin, 1933.
- [48] Marzke, R. F. and Wheeler, J. A., "Gravitation as Geometry—I: The Geometry of Space-Time and the Geometrodynamical Standard Meter," in *Gravitation and Relativity* (eds. H. Chen and W. F. Hoffman), W. A. Benjamin, New York, 1964.
- [49] Massey, G. J., *The Philosophy of Space*, unpublished doctoral dissertation, Princeton University, 1963.
- [50] Massey, G. J., "Toward a Clarification of Grünbaum's Concept of Intrinsic Metric," *Philosophy of Science*, vol. 36 (1969), pp. 331–345.
- [51] Massey, G. J., *Understanding Symbolic Logic*, Harper and Row, New York, 1970.
- [52] Massey, G. J., "Is 'Congruence' a Peculiar Predicate?" in *Boston Studies in the Philosophy of Science* (Proceedings of the 1970 [2nd] Biennial Congress of the Philosophy of Science Association), (eds. R. C. Buck, and R. S. Cohen), to be published by Reidel, Holland.
- [53] Menger, K., *Dimensionstheorie*, B. G. Teubner, Leipzig, 1928.
- [54] Möller, C., *The Theory of Relativity*, Oxford University Press, Oxford, 1952.
- [55] Newton, I., *Principia* (ed. F. Cajori), University of California Press, Berkeley, 1947.
- [56] Putnam, H., "An Examination of Grünbaum's Philosophy of Geometry," in *Philosophy of Science*, The Delaware Seminar, vol. 2 (ed. B. Baumrin), Interscience, New York, 1963, pp. 205–255.
- [57] Reichenbach, H., *Philosophie der Raum-Zeit-Lehre*, Walter de Gruyter and Co., Berlin, 1928.
- [58] Reichenbach, H., *The Philosophy of Space and Time*, Dover Publications, New York, 1958.
- [59] Reichenbach, H., *Axiomatization of the Theory of Relativity*, University of California Press, Berkeley, 1969.
- [60] Riemann, B., *Gesammelte Mathematische Werke und Wissenschaftlicher Nachlass*, 2nd ed. (ed. H. Weber), Dover Publications, New York, 1953.
- [61] Riemann, B., "On the Hypotheses Which Lie at the Foundations of Geometry," in *A Source Book in Mathematics*, vol. II (ed. D. E. Smith), Dover Publications, New York, 1959, pp. 411–425.
- [62] Russell, B., "Sur les Axiomes de la Géométrie," *Revue de Métaphysique et de Morale*, vol. VII (1899), pp. 684–707.
- [63] Russell, B., *The Foundations of Geometry*, Dover Publications, New York, 1956.
- [64] Schrödinger, E., *Expanding Universes*, Cambridge University Press, Cambridge, 1956.
- [65] Swinburne, R., Review of A. Grünbaum's *Geometry and Chronometry in Philosophical Perspective* in *The British Journal for the Philosophy of Science*, vol. 21 (1970) pp. 308–311.
- [66] Sygne, J. L., *Relativity: The Special Theory*, North-Holland Publishing, Amsterdam, 1956.
- [67] Tarski, A., "What is Elementary Geometry?," in *The Axiomatic Method* (eds. L. Henkin, P. Suppes and A. Tarski), North-Holland Publishing, Amsterdam, 1959.
- [68] van Fraassen, B., "On Massey's Explication of Grünbaum's Conception of Metric," *Philosophy of Science*, vol. 36 (1969), pp. 346–353.
- [69] Veblen, O., "The Foundations of Geometry," in *Modern Mathematics* (ed. J. W. A. Young), Dover Publications, New York, 1955, pp. 3–51.
- [70] Weatherburn, C. E., *Riemannian Geometry and the Tensor Calculus*, Cambridge University Press, Cambridge, Mass., 1957.
- [71] Weyl, H., *Space-Time-Matter*, Dover Publications, New York, 1950.
- [72] Wheeler, J. A., "Gravitation as Geometry—II," in *Gravitation and Relativity* (eds. H. Chen and W. F. Hoffman), W. A. Benjamin, New York, 1964.
- [73] Wheeler, J. A., *Geometrodynamics*, Academic Press, New York, 1962.
- [74] Wheeler, J. A., *Einstein's Vision*, Springer-Verlag, Berlin, 1968.
- [75] Wheeler, J. A., "Particles and Geometry," in *Relativity* (eds. M. Carmeli, S. I. Fickler, and L. Witten), Plenum Press, New York, 1970.

* Parts I–III of the present volume.

** Chapter 17 of the present volume.

‡ Chapter 20 of the present volume.

CAN WE ASCERTAIN THE FALSITY OF A SCIENTIFIC HYPOTHESIS?

See Append
§§48 and 49

1. INTRODUCTION

Can we ascertain the falsity of a given scientific hypothesis H ? Alternatively, could we ascertain the truth of H ? One tradition answers these questions *asymmetrically* as follows: Alas, unfavorable results furnished by just one kind of experiment suffice to guarantee the falsity of an otherwise highly successful hypothesis. And would that *favorable experimental* findings had a comparable capability of establishing the *truth* of a hypothesis! Thus, the scientist is held to be laboring under a discouraging handicap in his quest to glean nature's secrets. His most triumphant theories are never safe from refutation by potential contrary evidence. Hence, none of his hypotheses can ever be known to be true with certainty. But if even a small amount of contrary evidence does materialize, then the most celebrated of hypotheses is indeed known to be false.

On this view, the asymmetry between provability and *disprovability* arises from two simple facts of elementary deductive logic. If an observational consequence O of a hypothesis H turns out to be true, then the truth

of H does *not* follow deductively. The reason is that the truth of H is *not* necessary for the truth of O but only sufficient. To infer the truth of H deductively here would be to commit the so-called fallacy of affirming the consequent. Yet, if alternatively the observational findings do *not* accord with O, then the falsity of O does entail the falsity of H. For if H were true, *all* of its observational consequences, including O, would have to be true. The deduction of the falsity of H in this case is held to be valid by conforming to the so-called *modus tollens* form of valid inference.

This asymmetrical account of the verifiability vis-à-vis the falsifiability of a hypothesis leads to the claim that there are *CRUCIAL experiments* in science. For suppose that two rival hypotheses H_1 and H_2 each explain a given large set of observational facts. But suppose further that these hypotheses entail incompatible observational results O_1 and O_2 respectively as outcomes of a given kind of experiment C. And let the performance of C (on one or more occasions) issue in the observation O_1 . In that case, we are told, C is a crucial experiment because it decisively eliminates H_2 from the scientific arena as conclusively refuted by O_1 . Therefore, the experiment C confirms H_1 vis-à-vis H_2 .

But this confirmation of H_1 vis-à-vis H_2 is not at all tantamount to a conclusive verification of H_1 per se. After all, the outcome O_1 of C cannot render H_1 immune to refutation by potential contrary evidence O_3 . Indeed, the favorable outcome of C cannot assure the success of H_1 in a potential confrontation with another rival hypothesis H_3 in some other crucial experiment C'. In short, according to the asymmetrical account, there are crucially falsifying experiments but not crucially verifying ones.

The many treatises and textbooks of science which characterize certain experiments as crucial attest the influence of this account. For example, the experiment which Galileo allegedly performed from the Leaning Tower of Pisa is often presented as a decisive refutation of Aristotle's hypothesis that the acceleration of a heavy falling body is greater than that of a lighter one. And in a recent encyclopedia article, the renowned geneticist G. W. Beadle credits Louis Pasteur's experiments of 1862 with being crucial. Says he: "Louis Pasteur completely *disproved* the theory of spontaneous generation. He showed that bacteria would not grow in materials which were sterilized" (my italics).¹

But if Pasteur's experiments of 1862 did indeed constitute a complete disproof of the hypothesis of spontaneous generation of life, as claimed here by Beadle, then one wonders at once how that hypothesis could have been rehabilitated by A. I. Oparin in 1938 and further by H. Urey in 1952 to the following effect: life on earth originated by spontaneous generation under favorable conditions prevailing sometime between 4.5 billion years ago and the time of the earliest fossil evidence 2.7 billion years ago. Indeed, the conception of crucial experiment exemplified by Beadle's overstatement is an oversimplification of the falsifiability of a scientific hypothesis.² To show this, I shall first consider briefly the logic underlying the purported disproof

¹ G. W. Beadle, "Spontaneous Generation," *World Book Encyclopedia* (Toronto: Field Enterprises Educational Corp., 1967), vol. 17, p. 626.

² In his *Induction and Intuition in Scientific Thought* [American Philosophical Society Memoir, vol. 75, Philadelphia, 1969], Peter B. Medawar gave a telling illustration from *cancer research* (p. 54, n. 44) of the logical pitfalls which can jeopardize the reasoning underlying purported crucial experiments

of the occurrence of spontaneous generation of life. And then I shall examine an important episode in the history of recent astronomy culminating in the claim that R. H. Dicke's observations of the flattening of the sun disprove Einstein's general theory of relativity.

2. PURPORTED DISPROOFS OF HYPOTHESES IN BIOLOGY AND ASTRONOMY

1. *The Biological Hypothesis of Spontaneous Generation*

The spontaneous generation of life is sometimes called "abiogenesis." For the sake of brevity, I shall refer to the hypothesis of spontaneous generation as the "SG-hypothesis" or simply as "SG."

Presumably, SG states that living systems did and indeed would develop after an unspecified time interval under unspecified kinds of conditions which actually prevailed on the earth sometime during the past. But it plainly does *not* follow from *this* formulation of SG *alone* that *bacteria* would grow even once—let alone every time—in sterilized materials under the conditions of Pasteur's experiments during the time periods which he allotted to them. What then is the probative force of Pasteur's negative results as a basis for refuting SG? Let us suppose, just for argument's sake, that Pasteur's

pertaining to transformations of cells into the cancerous state. Karl Popper's influential version of the asymmetrical account of verifiability and falsifiability has very recently been criticized by B. Juhos, "Die methodologische Symmetrie von Verifikation und Falsifikation," *Zeitschrift für allgemeine Wissenschaftstheorie* I (1970): 41–70, especially 52ff.

own failure to obtain bacteria had established conclusively that there is no instance at all in which bacteria would grow spontaneously in his kind of experiment. Even then his experimental results could *not* have refuted the very general SG-hypothesis via *modus tollens*. The obvious reason is that a theory can be refuted by Pasteur's negative results *only* if that theory entails a positive or contrary outcome for his kind of experiment, which SG alone does not.

The entailment of a positive outcome could be assured by conjoining SG with a suitable auxiliary hypothesis *A* as follows: The auxiliary hypothesis *A* states that the time interval during which abiogenesis would occur need only be a small fraction of a man's life span, and furthermore *A* confers specificity on SG by supplying the kinds of physical conditions relevant to Pasteur's experiment, such as the presence of an oxidizing atmosphere.

Hence what Pasteur's negative results refute is *not* SG itself, but only the *conjunction* of SG with *A*. Indeed, SG can be upheld in the face of his findings by blaming the false prediction of a positive outcome of his experiments on the falsity of *A*. And *A* can therefore be replaced by a suitably different auxiliary hypothesis *A'* to confer specificity onto SG. No wonder, therefore, that Oparin, Urey, and others later felt free to postulate the truth of SG as part of a speculative, augmented theory. Their kind of theory asserted very roughly the following: When the earth was first formed, it had a reducing atmosphere of methane, ammonia, water, and hydrogen. Only at a later stage did photochemical splitting of water issue in an oxidizing atmosphere of carbon dioxide, nitrogen and oxygen. The

action of electric discharges or of ultra-violet light on a mixture of methane, ammonia, water, and hydrogen yields simple organic compounds such as amino acids and urea, as shown by work done since 1953.³ The first living organism originated by a series of non-biological steps from simple organic compounds which reacted to form structures of ever greater complexity until producing a structure that qualifies as living.

Commenting on this particular augmentation of SG, Beadle says that "no one has ever demonstrated or proved this development under satisfactorily controlled conditions."⁴ But as long as future evidence may still sustain these speculations, as Beadle seems to allow here, then it is clearly too strong to claim that "Louis Pasteur completely disproved the theory of spontaneous generation."

2. *The General Theory of Relativity*

Let us now turn to the evaluation of a recent purported experimental disproof of the general theory of relativity. To set the stage for this appraisal, let us first

³ S. L. Miller, "Life, Origin of," in *The International Encyclopedia of Science*, ed. James R. Newman (New York: Thomas Nelson & Sons, 1965), vol. II, p. 622. A. I. Oparin, *Genesis and Evolutionary Development of Life*, translated from the 1966 Russian edition (Moscow: Meditsina Publishing House) by Eleanor Maas (New York: Academic Press, 1968); also available as NASA Technical Translation TT F-88 (Washington: NASA, 1968). See also J. D. Bernal, *The Origin of Life* (Cleveland: World, 1967); D. H. Kenyon and G. Steinman, *Biological Predestination* (New York: McGraw-Hill, 1969); and P. Handler, ed., *Biology and the Future of Man* (New York: Oxford University Press, 1970), Chap. 5, "The Origins of Life."

⁴ Beadle, "Spontaneous Generation."

have a very brief look at the logic of the threat posed to Newton's theory of motion and gravitation by two sets of observational findings. First, there was a discrepancy between the *observed* motion of the planet Uranus and its motion as *calculated* theoretically about 1820. This calculation was obtained from Newton's theory via Laplace's determination of the perturbations superposed by Jupiter and Saturn on the sun's gravitational attraction. The second threat to Newton's theory arose from the magnitude of the observed slow precession of Mercury's perihelion or of its longer orbital axis in the direction of its motion. The observed precession yielded a troublesome excess of about 43 seconds of arc per century. In Newton's theory, the orbit of a planet due to the sun's gravitational attraction alone would be a closed ellipse with the sun at one focus. Hence the Newtonian will seek to account for the slow rotation of Mercury's ellipse by means of perturbations emanating from the other planets, notably from Venus, which is the closest, and from Jupiter, which is the most massive. But such a Newtonian calculation fails to account for a residuum of 43 seconds of arc per century from the observed value of the precession. Let us see how Newton's theory attempted to deal with the challenge of these two sets of observations.

In the case of Uranus, the theory from which its orbit was calculated in about 1820 included the hypothesis *H* of Newton's laws and the auxiliary assumption *A* that the relevant perturbations emanated *entirely* from Jupiter and Saturn. Historically, the empirical incorrectness of the calculated orbit of Uranus was blamed *not* on the falsity of the Newtonian *H* but rather on the falsity of *A*. Specifically, Adams and Leverrier each

theoretically postulated the previously unsuspected existence of the planet Neptune, thereby introducing a contrary auxiliary hypothesis A'. It was the first case of inferring a perturbing planet from the observed perturbations rather than conversely.⁵ Ironically, Lalande had unwittingly observed Neptune as early as May 1795. Having assumed it to be a fixed star, he attributed the discordance between its apparent positions two days apart to observational errors and discounted at least one of these two observations.⁶ It is interesting to note what probative significance was attributed to this detection of Neptune through its action on Uranus before Neptune's presence was corroborated observationally. The American astronomer, Simon Newcomb (1835-1909), hailed it as "a striking example of the precision reached by the theory of the celestial motions."⁷ The theory to which Newcomb attributes this virtue here includes among its component premises Newton's H and an A' which asserts the existence of *eight* of the now known nine planets.⁸

But this very combination of H and A' yielded an observationally *incorrect* value for the precession of the perihelion of Mercury. The astronomers then made clear how difficult it is to disprove observationally once and for all any one component hypothesis of a total theoretical system. For they noted that the observed

⁵ For details, see A. Pannekoek, *A History of Astronomy* (New York: Interscience Publishers, 1961), pp. 359-63.

⁶ "Neptune," *The Encyclopædia Britannica*, 14th ed., 1929, vol. 16, p. 228.

⁷ *Ibid.*, p. 226.

⁸ Pluto was not inferred from the perturbations in Neptune's own orbit until 1931.

greater value does not enable us to pinpoint the blame as between the components H and A' of the total celestial mechanics. Thus, assuming that only *one* of the two components H and A' is false, Leverrier considered a new auxiliary hypothesis A'' which postulated a group of planets revolving inside the orbit of Mercury to account for the discrepancy. And *at first*, it even seemed that the existence of such intra-Mercurial bodies was confirmed by occasional reports of dark objects seen in transit across the sun's face. But numerous photographs failed to bear out Leverrier's A''.⁹ Hence other astronomers allowed that Newton's H might be false. They modified Newton's law of gravitation by assuming that its exponent is slightly greater than 2 in the amount of $1/6,000,000$.¹⁰ But this attempt to modify Newton's theory gave way to the general theory of relativity. In Einstein's theory the orbit of a planet of negligibly small mass *subject solely to the sun's gravitational field* is *not* a perfectly closed ellipse about the sun, as it is in Newton's theory. Instead, the orbit of such a planet is a slowly rotating ellipse.¹¹ For Mercury, which is the planet closest to the sun, the amount of the perihelion precession thus yielded by general relativity *without* any allowance for the effects of any other planets is about 43 seconds of arc per century. This is in satisfactory agreement with the residual portion of the ob-

⁹ Cf. "Mercury, Motion of Perihelion," *Encyclopædia Britannica*, 14th ed., 1929, vol. 15, p. 269.

¹⁰ Pannekoek, *A History of Astronomy*, p. 363.

¹¹ See A. Einstein, "The Foundations of the General Theory of Relativity," in *The Principle of Relativity, a Collection of Original Memoirs* (New York: Dover Publications, 1952), pp. 163-64; and R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford: Oxford University Press, 1934), pp. 208-9.

served value that had been left unexplained by the perturbations from other planets in Newton's theory.¹² This agreement prompted the Secretary of the Royal Astronomical Society to write in 1929 that general relativity had "satisfactorily cleared up" the difficulty left unresolved by Newtonian theory.¹³ Note that the Newtonian theory deduces a perihelion precession *only* because it takes cognizance of the effects of other planets. By contrast, the relativity theory obtains a *part* of the total precession from the gravitational effect of the sun alone. But the relativistic calculation treated the sun's field as spherically symmetrical. Thereby it neglected any gravitational effects on Mercury arising from such oblateness of the sun as is due to the centrifugal effects of its own rotation.¹⁴

Suppose that the entire sun rotates with a constant angular velocity equal to the average of that observed on different parts of its surface, i.e., with a period of about 25 days. In that case, centrifugal effects yield an amount of oblateness of the sun's mass distribution that can be neglected with impunity apropos of Mercury's perihelion. But in a 1967 paper, R. H. Dicke and H. M. Goldenberg reported an oblateness of the distribution

¹² See V. L. Ginzburg, "Experimental Verifications of the General Theory of Relativity," in *Recent Developments in General Relativity* (New York: Pergamon Press, 1962), pp. 58-60; and R. H. Dicke, *The Theoretical Significance of Experimental Relativity* (New York: Gordon and Breach, 1964), pp. 27-28.

¹³ Cf. "Mercury, Motion of Perihelion," *Encyclopædia Britannica*, p. 269.

¹⁴ H. P. Robertson and T. W. Noonan, *Relativity and Cosmology* (Philadelphia: W. B. Saunders Co., 1968), p. 239; and L. I. Schiff, "Gravitation and Relativity," in D. L. Arm, ed., *Journeys in Science* (Albuquerque: University of New Mexico Press, 1967), pp. 153 and 156.

of the sun's visible brightness corresponding to more than four times the above neglected amount.¹⁵ They interpret this *visual* oblateness as being associated with a corresponding degree of oblateness in the sun's *mass* distribution! And in order to reconcile this amount of solar equatorial bulge with the 25-day period of rotation of the sun's surface, they assume that a core in the sun rotates as rapidly as once in a day or two.

But, alas, according to both Newtonian physics and relativity theory, an equatorial bulge in the sun's mass distribution of the magnitude claimed by Dicke and Goldenberg alone makes for a precession of Mercury's perihelion in the amount of 3.4 seconds of arc per century. (Here the Newtonian calculation is an excellent relativistic approximation!)¹⁶ This amount of rotation of Mercury's ellipse must be added to the relativistic value of 43 seconds, which prevails in the absence of any oblateness. Hence the conjunction of Einstein's equations with the assertion of an oblate mass distribution in the amount assumed by Dicke entails the value of 46.4 seconds of arc per century. Thus, if the sun's mass distribution does deviate from spherical symmetry in the amount claimed by Dicke, and *if*, furthermore, the observations of Mercury's orbit can be taken as proving that the actual value of the unexplained residuum is significantly different from 46.4 seconds, then Einstein's general theory of relativity gives rise to a false conclusion. And, in that case, relativity theory has been refuted in Dicke's laboratory!

¹⁵ R. H. Dicke and H. M. Goldenberg, *Physical Review Letters* 9 (1967), p. 313.

¹⁶ See Schiff, "Gravitation and Relativity," p. 156.

It is evident, therefore, that Dicke's disproof of general relativity is predicated, among other things, on the following important assumption: the observed amount of *optical* oblateness of the sun *confirms* it to be *true* that the sun's mass distribution deviates correspondingly from spherical symmetry, i.e., establishes that the sun is also oblate *gravitationally*. To be sure, if gravitational oblateness were present, then it would serve to explain the observed optical flattening. But the existence of optical flattening does *not* entail, and hence cannot guarantee deductively, that there is gravitational oblateness. The inference from optical to gravitational oblateness is *inductive* and uncertain rather than deductive and certain. Indeed, it was pointed out in 1967 that there are at least two possible explanations of the sun's apparent equatorial bulge, neither of which involves any appreciable gravitational oblateness at all. And moreover, it has been argued that Dicke's very high differential rotation rate between the sun's core and surface is too unstable to endure for any length of time.¹⁷

No wonder, therefore, that there is disagreement in the scientific community as to whether Dicke's findings do disprove Einstein's general relativity. Thus, L. I. Schiff writes:

At the present time we cannot conclude that the observed solar oblateness invalidates general relativity theory. On the contrary, in view of . . . the difficulties inherent in Dicke and Goldenberg's interpretation of their observations, it seems most reasonable to assume for the present that Einstein's theory is correct.¹⁸

¹⁷ *Ibid.*, p. 157. See also note 76 at the end of this essay.

¹⁸ *Ibid.*

This episode from the history of recent and current astronomy *suggests* the following quite *fundamental epistemological conclusion*:

THE IRREMEDEIABLE INCONCLUSIVENESS OF THE VERIFICATION OF AN AUXILIARY COMPONENT OF A TOTAL THEORY BY ITS SUPPORTING EVIDENCE IMPOSES A CORRESPONDING LIMITATION ON THE DEDUCIBILITY OF THE CATEGORICAL FALSITY OF THE MAIN COMPONENT OF THE TOTAL THEORY VIA OTHER EVIDENCE ADVERSE TO THE TOTAL THEORY.

In other words, given that the conjunction $H \cdot A$ entails a false prediction, then this fact alone does *not* enable us to deduce the falsity of H itself, unless A is known to be true. But since the truth of A cannot be *guaranteed* by its supporting evidence, the falsity of H is correspondingly uncertain. A two-stage schematization of the logic of our astronomical example will lend greater specificity to this suggested epistemological moral.

- | | |
|-----|---|
| I. | $\frac{\text{GTR} \cdot A \text{ (which includes DP)}}{\downarrow \text{ (entails)}}$ $C \text{ (46.4 seconds of arc/century)}$ $\text{But } O \text{ (43 seconds of arc/century),}$ $\text{so that } O \rightarrow \sim C$ $\therefore \sim (\text{GTR} \cdot A)$ <hr style="width: 100%;"/> |
| II. | $\sim \text{GTR} \vee \sim A$ $\text{But also } A$ <hr style="width: 100%;"/> $\therefore \sim \text{GTR}$ |

Let "GTR" be an abbreviation for "general theory of relativity," and let Dicke's postulate of *gravitational* oblateness be called "Dicke's Postulate" or "DP." Further-

are each false, as could readily happen to be the case. Then the presumed experimental falsity of C_2 fully *allows* the following state of affairs: Although it is considered decisively disproven in a crucial experiment, H_2 is actually true. And the conjunction of H_2 with A_2 entailed a false conclusion C_2 *only* because A_2 is false! Moreover, H_1 is actually false, although the conjunction of the false H_1 with the likewise false A_1 does entail the true consequence C_1 . Under these circumstances, the *false* H_1 will emerge triumphant in the scientific community over the true H_2 . And those who regard the experiment as crucial will mistakenly discard H_2 as *beyond any possible rehabilitation*. In this case, the supposed crucial experiment is surely quite counterproductive. Needless to say, the experiment might alternatively have been recognized as failing to *decide between* H_1 and H_2 if it had yielded a result incompatible with *both* C_1 and C_2 .

Thus, the logical situations encountered in our historical examples exhibit the inadequacy of the received textbook account, according to which a component hypothesis can be refuted once and for all. The great French historian of science, Pierre Duhem, pioneered in making us aware of other such episodes in the earlier history of science. And he bequeathed a very influential philosophical legacy to us on the issue of the falsifiability of a scientific hypothesis. In pursuing our central concern with this issue, the remainder of this essay will therefore address itself to Duhem's philosophical legacy. Let us recall the fundamental epistemological moral *suggested* by our example from recent astronomy. Then we can state our central problem more precisely as follows: Can we *ever* justifiably reject a component

hypothesis of a theory as irrevocably falsified by observation?

3. IS IT NEVER POSSIBLE TO FALSIFY A HYPOTHESIS IRREVOCABLY?

In his book *The Aim and Structure of Physical Theory*, Duhem denied the feasibility of crucial experiments in physics. Said he:

. . . the physicist can never subject an isolated hypothesis to experimental test but only a whole group of hypotheses; when the experiment is in disagreement with his predictions, what he learns is that at least one of the hypotheses constituting this group is unacceptable and ought to be modified; *but the experiment does not designate which one should be changed* (my italics).¹⁹

Duhem illustrates and elaborates this contention by means of examples from the history of optics. And in each of these cases, he maintains that "If physicists had attached some value to this task,"²⁰ any one component hypothesis of optical theory such as the corpuscular hypothesis (or so-called emission hypothesis) could have been preserved in the face of seemingly refuting experimental results such as those yielded by Foucault's experiment. According to Duhem, this continued espousal of the component hypothesis could be justified by "shifting the weight of the experimental contradiction to some other proposition of the commonly accepted optics."²¹ Here Duhem is maintaining that the

¹⁹ Pierre Duhem, *The Aim and Structure of Physical Theory* (Princeton: Princeton University Press, 1954), p. 187.

²⁰ *Ibid.*

²¹ *Ibid.*, p. 186.

refutation of a component hypothesis *H* is at least usually no more certain than its verification could be.

In terms of the notation *H* and *A* which we have been using, Duhem is telling us that we could blame an experimentally false consequence *C* of the total optical theory *T* on the falsity of *A* while upholding *H*. In making this claim, Duhem is quite clear that the falsity of *A* no more *follows* from the experimental falsity of *C* than does the falsity of *H*. But his point is that this fact does not logically prevent us from postulating that *A* is false while *H* is true. And he is telling us that altogether the pertinent empirical facts *allow* us to reject *A* as false. Hence we are under no deductive logical constraint to infer the falsity of *H*. In questioning Dicke's purported refutation of the GTR, I gave sanction to this Duhemian contention in this instance precisely on the grounds that Dicke's auxiliary *A* is *not* known to be true with certainty.

In contemporary philosophy of science, a *generalized* version of Duhem's thesis with its ramifications has been attributed to Duhem and has been highly influential. I shall refer to this elaboration of Duhem's philosophical legacy in present-day philosophy of science as "the D-thesis." But in doing so, my concern is with the philosophical credentials of this legacy, not with whether this attribution to Duhem himself can be uniquely sustained exegetically as against rival interpretations given by Duhem scholars. But I should remark that at least one such scholar, L. Laudan, has cited textual evidence which casts some doubt on this attribution.²² The pres-

²² Laurens Laudan, "On the Impossibility of Crucial Falsifying Experiments: Grünbaum on 'The Duhemian Argument'," *Philosophy of Science* 32 (1965): 295–99.

ent philosophical appraisal is intended to supersede some parts of my earlier published critique of the D-thesis. In his Pittsburgh Doctoral Dissertation, Philip Quinn has pointed out that the version of the D-thesis with which I am concerned can be usefully stated in the form of two subtheses D1 and D2, and has argued that Laudan's attributional doubts are warranted only with respect to D2 but not with respect to D1.

The two subtheses are:

D1. No constituent hypothesis H of a wider theory can *ever* be sufficiently isolated from some set or other of auxiliary assumptions so as to be separately falsifiable observationally. H is here understood to be a constituent of a wider theory in the sense that no observational consequence can be deduced from H alone.

It is a corollary of this subthesis that *no* such hypothesis H *ever* lends itself to a crucially falsifying experiment any more than it does to a crucially verifying test.

D2. In order to state the second subthesis D2, we let T be a theory of *any* domain of empirical knowledge, and we let H be *any* of its component subhypotheses, while A is the collection of the remainder of its subhypotheses. Also, we assume that the observationally testable consequence O entailed by the conjunction $H \cdot A$ is taken to be empirically false, because the observed findings are taken to have yielded a result O' *incompatible* with O. Then D2 asserts the following: For all potential empirical findings O' of this kind, there exists at least one suitably revised set of auxiliary assumptions A' such that the conjunction of H with A' *can be held to be true and explains O'*. Thus D2 claims that H can be held to be true *and* can be used to explain O' no matter what O' turns out to be, i.e., *come what may*.

Note that if D2 did not assert that A' can be held to be true in the face of the evidence no less than H, then H could not be claimed to explain O' via A'. For premises which are already *known* to be false are not scientifically acceptable as bases for explanation.²³ Hence the part of D2 which asserts that H and A' can each be held to be *true* presupposes that either they could not be separately falsified or that neither of them has been separately falsified.

In my prior writings on the D-thesis, I made three main claims concerning it:

1. There are quite trivial senses in which D1 and D2 are uninterestingly true and in which no one would wish to contest them.²⁴
2. In its non-trivial form, D2 has not been demonstrated.²⁵
3. D1 is false, as shown by counterexamples from physical geometry.²⁶

²³ For a discussion of the epistemic requirement that explanatory premises must *not* be known to be false, see Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace and World, 1961), pp. 42–43.

²⁴ A. Grünbaum, "The Falsifiability of a Component of a Theoretical System," in *Mind, Matter, and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*, P. K. Feyerabend and G. Maxwell, eds. (Minneapolis: University of Minnesota Press, 1966), pp. 276–80.

²⁵ *Ibid.*, pp. 280–81.

²⁶ *Ibid.*, pp. 283–95; and A. Grünbaum, *Geometry and Chronometry in Philosophical Perspective* (Minneapolis: University of Minnesota Press, 1968), Chap. III, pp. 341–51. In Chap. III, §9.2, pp. 351–69 of the latter book, I present a counterexample to H. Putnam's *particular* geometrical version of D2. For a brief summary of Putnam's version, see footnote 69.

Since then, Gerald Massey has called my attention to yet another defect of D2 which any proponent of that thesis would presumably endeavor to remedy. Massey has pointed out that, as it stands, D2 attributes *universal* explanatory relevance and power to any one component hypothesis H. For let O^* be *any* observationally testable statement *whatever* which is compatible with $\sim O$, while O' is the conjunction $\sim O \cdot O^*$. Assume that $(H \cdot A) \rightarrow O$. Then D2 asserts the existence of an auxiliary A' such that the theoretical conjunction $H \cdot A'$ explains the putative observational finding $\sim O \cdot O^*$. As Joseph Camp has suggested, the proponent of D2 might reply that A' itself may potentially explain O^* *without* H, even though H is essential for explaining $\sim O$ via A' . But the advocate of D2 has no guarantee that he can circumvent the difficulty in this way.

Instead, he might perhaps wish to require that O' must pertain to *the same kind of phenomena* as O , thereby ruling out "extraneous" findings O^* . Yet even if he can articulate such a restriction or provide a viable alternative to it, there is the following further difficulty noted by Massey: D2 gratuitously asserts the existence of a *deductive* explanation for any event whatever. This existential claim is gratuitous. For there *may* be individual occurrences (in the domain of quantum phenomena or elsewhere) which cannot be explained deductively, because of the irreducibly statistical character of its pertinent laws.

Our governing concern here is the question: "Is there *any* component hypothesis H whatever whose falsity we can ascertain?" I shall try to answer this question by giving reasons for now *qualifying* my erstwhile charge that D1 is false. Hence I shall modify the *third* of my

earlier contentions about the D-thesis. But before doing so, I must give my reasons for *not* also retracting either of the first two of these contentions in response to the critical literature which they have elicited. These reasons will occupy a number of the pages that follow. The first of my earlier claims was that D1 and D2 are each true in trivial senses which are respectively exemplified by the following two examples, which I had given:²⁷

i. Suppose that someone were to assert the presumably false empirical hypothesis H that "Ordinary buttermilk is highly toxic to humans." Then the English sentence expressing this hypothesis could be "saved from refutation" in the face of the observed wholesomeness of ordinary buttermilk by making the following change in the theoretical system constituted by the hypothesis: changing the rules of English usage so that the intension of the term "ordinary buttermilk" is that of the term "arsenic" as customarily understood. In this Pickwickian sense, D1 could be sustained in the case of this particular H.

ii. In an endeavor to justify D2, let someone propose the use of an A' which is itself of the form

$$\sim H \vee O'.$$

In that case, it is certainly true in standard systems of logic that

$$(H \cdot A') \rightarrow O'.$$

Let us now see by reference to these two examples why I regard them as exemplifications of trivially true versions of D1 and D2 respectively.

²⁷ Grünbaum, "The Falsifiability of a Component of a Theoretical System," pp. 277-78.

i. *D1*

Here *H* was the hypothesis that "ordinary buttermilk is highly toxic to humans." When the proponent of *D1* was challenged to save this *H* from refutation, what he did "save" was *not* the proposition *H* but the *sentence* *H* expressing it, as reinterpreted in the following respect: Only the term "ordinary buttermilk" was given a new semantical usage, and *no constraint was imposed on its new usage other than that the ensuing reinterpretation turn the sentence H into a true proposition.*

If one does countenance such *unbridled* semantical instability of some of the theoretical language in which *H* is stated, then one can indeed thereby uphold *D1* in the form of Quine's epigram: "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system."²⁸ But, in that case, *D1* turns into a thoroughly unenlightening truism.

I took pains to point out, however, that the commitments of *D1* can also be trivially or uninterestingly fulfilled by semantical devices far more sophisticated and restrictive than the unbridled reinterpretation of some of the vocabulary in *H*. As an example of such a trivial fulfillment of *D1* by devices whose feasibility is itself not at all trivial in other respects, I cited the following:

. . . suppose we had two particular substances I_1 and I_2 which are isomeric with each other. That is to say, these substances are composed of the same elements in the same proportions and with the same molecular weight

²⁸ W. V. O. Quine, *From a Logical Point of View* (revised ed.), (Cambridge, Mass.: Harvard University Press, 1961), pp. 43 and 41n.

but the arrangement of the atoms within the molecule is different. Suppose further that I_1 is not at all toxic while I_2 is highly toxic, as in the case of two isomers of trinitrobenzene. [At this point, I gave a footnote citation of 1, 3, 5-trinitrobenzene and of 1, 2, 4-trinitrobenzene respectively.] Then if we were to call I_1 "aposteriorine" and asserted that "aposteriorine is highly toxic," this statement H could also be trivially saved from refutation in the face of the evidence of the wholesomeness of I_1 by the following device: only *partially* changing the meaning of "aposteriorine" so that its intension is the second, highly toxic isomer I_2 , thereby leaving the chemical "core meaning" of "aposteriorine" intact. To avoid misunderstanding of my charge of triviality, let me point out precisely what I regard as trivial here. The preservation of H from refutation in the face of the evidence by a *partial* change in the meaning of "aposteriorine" is trivial in the sense of being only a *trivial* fulfillment of *the expectations raised by the D-thesis* [D1]. But, in my view, *the possibility as such of preserving H by this particular kind of change in meaning* is not at all trivial. For this possibility as such reflects a fact about the world: the existence of isomeric substances of radically different degrees of toxicity (allergenicity)!²⁹

Mindful of this latter kind of example, I emphasized that a construal of D1 which allows itself to be sustained by this kind of alteration of the intension of "trinitrobenzene" is no less trivial *in the context of the expectations raised by the D-thesis* than one which rests its case on calling arsenic "buttermilk."³⁰ And hence I

²⁹ Grünbaum, "The Falsifiability of a Component of a Theoretical System," pp. 279–80.

³⁰ *Ibid.*, p. 279.

was prompted to conclude that "a necessary condition for the nontriviality of Duhem's thesis is that the theoretical language be semantically stable in the relevant respects [i.e., with respect to the vocabulary used in H]." ³¹

Mary Hesse took issue with this conclusion in her review of the essay in which I made these claims. There she wrote:

. . . it is not clear . . . that "semantic stability" is always required when a hypothesis is nontrivially saved in face of undermining evidence. The law of conservation of momentum is in a sense saved in relativistic mechanics, and yet the usage of "mass" is changed—it becomes a function of velocity instead of a constant property. But further argument along these lines is idle without more detailed analysis of what it is for a hypothesis to be the "same," and what is involved in "semantic stability." ³²

This criticism calls for several comments.

(a) Hesse calls for a "more detailed analysis of what it is for a hypothesis to be the 'same,' and what is involved in 'semantic stability.'" To this I say that the *primary* onus for providing that more detailed analysis falls on the shoulders of the Duhemian. For he wishes to claim that his thesis D1 is interestingly true. And we saw that if he is to make such a claim, he must surely *not* be satisfied with the mere retention of the *sentence* H in some interpretation or other.

Fortunately, both the proponent and the critic of D1 can avail themselves of an apparatus of distinctions pro-

³¹ *Ibid.*, p. 278.

³² Mary Hesse, *The British Journal for the Philosophy of Science*, 18 (1968): 334.

posed by Peter Achinstein to confer specificity on what it is for the vocabulary of H to remain semantically stable.

Achinstein introduces semantical categories for describing the various possible relationships between properties of X and the term "X." And he appeals to the normal, standard or typical scientific *use* of a term "X" at a given time as distinct from special uses.³³ He writes:

. . . I must introduce the concept of relevance and speak of a property as relevant for being an X. By this I mean that if an item is known to possess certain properties and lack others, the fact that the item possesses (or lacks) the property in question normally will count, at least to some extent, in favor of (or against) concluding that it is an X; and if it is known to possess or lack sufficiently many properties of certain sorts, the fact that the item possesses or lacks the property in question may justifiably be held to settle whether it is an X.³⁴ . . .

Two distinctions are now possible. The first is between positive and negative relevance. If the fact that an item has P tends to count more in favor of concluding that it is an X than the fact that it lacks P tends to count against it, P can be said to have more positive than negative relevance for X. The second distinction is between semantical and nonsemantical relevance and is applicable only to certain cases of relevance.

Suppose one is asked to justify the claim that the reddish metallic element of atomic number 29, which is a good conductor and melts at 1083°C, is copper. One reply is that such properties tend to count in and of them-

³³ Peter Achinstein, *Concepts of Science* (Baltimore: Johns Hopkins Press, 1968), pp. 3ff.

³⁴ *Ibid.*, p. 6.

selves, to some extent, toward classifying something as copper. By this I mean that an item is correctly classifiable as copper solely in virtue of having such properties; they are among the properties which constitute a final court of appeal when considering matters of classification; such properties are, one might say, intrinsically copper-making ones. Suppose, on the other hand, one is asked to justify the claim that the substance constituting about 10⁻⁴ per cent of the igneous rocks in the earth's crust, that is mined in Michigan, and that was used by the ancient Greeks, is copper. Among the possible replies is *not* that such properties tend to count in and of themselves, to some extent, toward something's being classifiable as copper—that is, it is not true that something is classifiable as copper solely in virtue of having such properties. These properties do not constitute a final court of appeal when considering matters of classification. They are not intrinsically copper-making ones. Rather, the possession of such properties (among others) counts in favor of classifying something as copper solely because it allows one to infer that the item possesses other properties such as being metallic and having the atomic number 29, properties that are intrinsically copper-making ones, in virtue of which it is classifiable as copper.³⁵ . . .

Suppose $P^1, \dots, P_n$ constitutes some set of relevant properties of X. If the properties in this set tend to count in and of themselves, to some extent, toward an item's being classifiable as an X, I shall speak of them as *semantically relevant* for X. If the possession of properties by an item tends to count toward an X-classification solely because it allows one to infer that the item possesses properties of the former sort, I shall speak of such properties as *nonsemantically relevant* for X. This

³⁵ *Ibid.*, pp. 7–8.

distinction is not meant to apply to all relevant properties of X, for there will be cases on or near the borderline not clearly classifiable in either way.³⁶

. . . in the latter part of the eighteenth century, with the systematic chemical nomenclature of Bergman and Lavoisier, the chemical composition of compounds began to be treated as semantically relevant; and . . . chemical composition . . . provided the basis for classification of compounds.

I have used the labels semantical and nonsemantical relevance because X's semantically relevant properties have something to do with the meaning or use of the term "X" in a way that X's nonsemantically relevant properties do not.³⁷ . . .

Suppose you learn the semantically relevant properties of items denoted by the term "X." Then you will know those properties a possession of which by actual and hypothetical substances in and of itself tends to count in favor of classifying them as ones to which the term "X" is applicable.³⁸ . . . Properties semantically relevant for X . . . include those that are logically necessary or sufficient. . . .

Consider a term "X" and the properties or conditions semantically relevant for X. It is perfectly possible that there be two different theories in which the term "X" is used, where the same set of semantically relevant properties of X (or conditions for X) are presupposed in each theory (even though other properties attributed to X by these theories, properties not semantically relevant for X, might be different). If so, the term "X" would not mean something different in each theory.³⁹

³⁶ *Ibid.*, pp. 8–9.

³⁷ *Ibid.*, p. 9.

³⁸ *Ibid.*, p. 35.

³⁹ *Ibid.*, p. 101.

Let us now employ Achinstein's distinctions to characterize a semantically stable use of the sentence *H*. I would say that one is engaged in a *semantically stable use* of the term "X," if and only if no changes are made in the membership of the set of properties which are semantically relevant to being an item denoted by the term "X." And similarly for the semantically stable use of the various terms " X_i " ($i = 1, 2, 3, \dots, n$) constituting the vocabulary of the sentence *H* in which a particular hypothesis (proposition) is expressed, even though these terms will, of course, not be confined to substance words or to the three major types of terms treated by Achinstein.⁴⁰ By the same token, if the entire sentence *H* is used in a semantically stable manner, then the hypothesis *H* has remained the same in the face of other changes in the total theory. Moreover, employing different terminology, I dealt with the concrete case of geometry and optics in earlier publications to point out the following: when the rejection of a certain presumed law of optics leads to a change in the membership of the properties *non*semantically relevant to a geometrical term "X," this change does *not* itself make for a semantic instability in the use of "X."⁴¹ I shall develop the latter point further below after discussing Hesse's objection.

In the case of my example of the two isomers of trinitrobenzene, it is clear from the very names of the two isomers that the molecular structure is semantically relevant or even logically necessary to being 1, 2, 4-

⁴⁰ *Ibid.*, p. 2.

⁴¹ See pp. 143-44 of the present volume; and A. Grünbaum, *Geometry and Chronometry* . . . , pp. 314-17.

Grünbaum, *Geometry and Chronometry* . . . , pp. 314-17.

trinitrobenzene as distinct from 1, 3, 5-trinitrobenzene or conversely. It will be recalled that presumably only the *former* of these two isomers is toxic. The Duhemian should be concerned to be able to uphold the hypothesis "1, 3, 5-trinitrobenzene is toxic" (H) in a semantically stable manner. Hence, if he is to succeed in doing so, his use of the word "1, 3, 5-trinitrobenzene" must leave all of the semantically relevant properties of 1, 3, 5-trinitrobenzene unchanged. And thus it would constitute only a trivial fulfillment of D1 in this case to adopt a *new use* of "1, 3, 5-trinitrobenzene" which suddenly confers *positive* semantic relevance on the particular properties constituting the molecular structure of 1, 2, 4-trinitrobenzene. I do say that upholding H by such a "meaning-switch" is trivial *for the purposes of D1*. But I do *not* thereby affirm at all the biochemical triviality of the fact which makes it possible to uphold H in this particular fashion. For I would be the first to grant that the existence of isomers of trinitrobenzene differing greatly in toxicity is biochemically significant.

(b) Hesse believes that Einstein's replacement of the Newtonian law of conservation of momentum by the relativistic one is a case in which "a hypothesis is non-trivially saved in face of undermining evidence" amid a violation of semantic stability. And she views this case as a counterexample to my claim that semantic stability is a necessary condition for the non-trivial *fulfillment of D1*. The extension of my remarks about the trinitrobenzene example to this case will now serve to show why I do not consider this objection cogent: it rests on a conflation of being non-trivial in some respect or other with being non-trivial vis-à-vis D1.

One postulational base of the special relativistic dy-

namics of particles combines *formal homologues* of the two principles of the conservation of mass and momentum with the kinematical Lorentz transformations.⁴² It is then shown that it is possible to satisfy the two formal conservation principles such that they are Lorentz-covariant, only on the assumption that the mass of a particle depends on its velocity. And the exact form of that velocity-dependence is derived via the requirement that the conservation laws go over into the classical laws for moderate velocities, i.e., that the relativistic mass m assume the value of the Newton mass m_0 for vanishing velocity.⁴³ This latter fact certainly makes it non-trivial and useful *for mechanics* to use the word "mass" in the case of both Newtonian and relativistic mass. And hence Einstein's *formal* retention of the conservation principles is certainly *not* an instance of an unbridled reinterpretation of them. Yet despite the interesting common feature of the term "mass," and the formal homology of the conservation principles, the two theories disagree here.

Using Achinstein's concept of mere *relevance*, we can say that the Newtonian and relativity theories disagree here as to the membership of the set of properties *relevant* to "mass." For it is clear that in Newton's theory, velocity-independence is positively *relevant*, in Achinstein's sense, to the term "mass." And it is likewise clear that velocity-dependence is similarly positively relevant in special relativistic dynamics. What is unclear is whether velocity independence and dependence respec-

⁴² See, for example, Tolman, *Relativity, Thermodynamics and Cosmology*, pp. 42-45.

⁴³ See P. G. Bergmann, *Introduction to the Theory of Relativity* (New York: Prentice-Hall, 1946), pp. 86-88.

tively are *semantically* or *nonsemantically* relevant. Thus, it is unclear whether relativistic dynamics modified the Newtonian conservation principles in semantically relevant ways or preserved them semantically while modifying only the properties *nonsemantically* relevant to "mass" and "velocity."

Let us determine the bearing of each of these two possibilities on the force of Hesse's purported counterexample to my claim that semantic stability is a necessary condition for the non-triviality of the D-thesis.

In the event of semantic relevance, semantic stability has been violated in a manner already illustrated by my trinitrobenzene example. For in that event, Newton's theory can be said to assert the conservation principles in an interpretation which assigns to the abstract word "mass" the value m_0 as its denotatum, whereas relativity theory rejects these principles as generally false in that interpretation. But a *non-trivial* fulfillment of D1 here would have required the retention of the hypothesis of conservation of momentum in an interpretation which preserves *all* of the properties semantically relevant to "mass," and to "velocity" for that matter. Since this requirement is not met on this construal, the *formal* relativistic retention of this conservation principle cannot qualify as a nontrivial fulfillment of D1. But the success of Einstein's particular semantical reinterpretation is, of course, highly illuminating in other respects.

On the other hand, suppose that relativistic dynamics has modified only the properties which are *non-semantically* relevant to "mass" and "velocity." In that event, its retention of the conservation hypothesis is indeed non-trivial vis-à-vis D1. But on that alternative construal, the relativistic introduction of a velocity-

dependence of mass would *not* involve a violation of semantic stability. And then this retention cannot furnish Hesse with a viable counterexample to my claim that semantic stability is necessary for non-triviality vis-à-vis D1.

Finally, suppose that here we are confronted with one of Achinstein's borderline cases such that the distinction between semantic and non-semantic relevance does not apply in the case of mass to the relevant property of velocity-dependence. In that case, we can characterize the transition from the Newtonian to the relativistic momentum conservation law with respect to "mass" as merely a repudiation of the Newtonian claim that velocity-independence is positively relevant in favor of asserting that velocity-dependence is thus relevant. And then this transition cannot be adduced as a retention of a hypothesis H which violates semantic stability while being non-trivial vis-à-vis D1.

Since there are borderline cases between semantic and non-semantic relevance, it is interesting that Duhem has recently been interpreted as denying that the distinction between them is ever scientifically pertinent. C. Giannoni reads Duhem as maintaining that actual scientific practice always accords the same semantic role to *all* relevant properties.⁴⁴ Thus, C. Giannoni writes:

Duhem is concerned primarily with quantities which are the subject of derivative measurement rather than fundamental measurement. . . . Quantities which are measured by other means than the means originally used in introducing the concept are derivative relative to this

⁴⁴ C. Giannoni, "Quine, Grünbaum, and The Duhemian Thesis," *Nous* 1 (1967): 288

method of measurement. For example, length can be fundamentally measured by using meter sticks, but it also can be measured by sending a light beam along the length and back and measuring the time which it takes to complete the round trip. We can then calculate the length by finding the product of the velocity of light (c) and the time. Such a method of measurement is dependent not only on the measurement of time as is the first type of derivative quantity, but also on the law of nature that the velocity of light is a constant equal to c .⁴⁵

Duhem further substantiates his view . . . by noting that when several methods are available for measuring a certain property, no one method is taken as the absolute criterion relative to which the other methods are derivative in the second sense of derivative measurement noted above [i.e., in Achinstein's sense of being nonsemantically relevant]. Each method is used as a check against the others.⁴⁶

If this thesis of *semantic parity* among *all* of the relevant properties were correct, then one might argue that a semantically stable use of the sentence H would simply not be possible in conjunction with each of two different auxiliary hypotheses A and A' of the following kind: A and A' contain *incompatible* law-like statements each of which pertains to one or more entities that are *also* designated by terms in the sentence H. In earlier

⁴⁵ *Ibid.*, pp. 286–87.

⁴⁶ *Ibid.*, p. 288. The example which Giannoni then goes on to cite from Duhem is one in which a very *weak* electric current was running through a battery and where, therefore, one or more indicators may fail to register its presence. In that case, the current will still be said to flow if one or another indicator yields a positive response. This particular case may well be a borderline one as between semantic and nonsemantic relevance.

publications, I used physical geometry and optics as a test case for these claims. There I considered the actual use of geometrical and optical language in standard relativity physics⁴⁷ as well as a hypothetical Duhemian modification of the optics of special relativity.⁴⁸ Let me now develop the import of these considerations. We shall see that it runs counter to the thesis of universal semantic *parity* and sustains the feasibility of semantic stability.

The Michelson-Morley experiment involves a comparison of the round-trip times of light along equal closed spatial paths in different directions of an inertial system. As is made explicit in the account of this experiment in standard treatises, the spatial distances along the arms of the interferometer *are measured by rigid rods*.⁴⁹ The issue which arose from this experiment was conceived to be the following: Are the round-trip times *along the equal spatial paths* in an inertial frame generally *unequal*, as claimed by the aether theory, or equal as asserted by special relativity? Hence both parties to the dispute agreed that, to within a certain accuracy, equal numerical verdicts furnished by rigid rods were positively semantically relevant to the geometrical relation term "spatially congruent" (as applied to line segments).⁵⁰ But the statement of the dis-

⁴⁷ Grünbaum, *Geometry and Chronometry* . . . , pp. 314–17.

⁴⁸ See pp. 143–44 of the present volume.

⁴⁹ E.g., in Bergmann, *Introduction to the Theory of Relativity*, pp. 23–26.

⁵⁰ Since the aether-theoretically expected time difference in the second order terms is only of the order of 10^{-15} second, allowance had to be made in practice for the absence of a corresponding accuracy in the measurement of the equality of the two arms.

pute shows that the *positive* relevance of the equality of the round-trip times of light to *spatial* congruence was made *contingent* on the particular *law of optics* which would be borne out by the Michelson-Morley experiment. Let X be the relational property of being congruent as obtaining among intervals of space. And let Y be the relational property of being traversed by light in equal round-trip times, as applied to spatial paths as well. Then we can say that the usage of "X" at the turn of the current century was such that the relation Y was only *nonsemantically* relevant to the relation term "X," whereas identical numerical findings of *rigid rods*—to be called "RR"—were held to be positively semantically relevant to it.

Relativity physics then added the following results pertinent to the kind of relevance which can be claimed for Y: (1) Contrary to the situation in inertial frames, in the so-called "time-orthogonal" *non-inertial* frames the round-trip times of light in different directions along spatially *equal* paths are generally unequal.⁵¹ (2) On a disk rotating uniformly in an inertial system I, there is generally an *inequality* among the transit times required by light departing from the same point to traverse a given closed polygonal path in opposite

This is made feasible by the fact that, on the aether theory, the effect of any discrepancy in the lengths of the two arms should *vary*, on account of the earth's motion, as the apparatus is rotated. For details, see Bergmann, *Introduction to the Theory of Relativity*, pp. 24–26, and J. Aharoni, *The Special Theory of Relativity* (Oxford: Oxford University Press, 1959), pp. 270–73. Indeed, slightly unequal arms are needed to produce neat interference fringes.

⁵¹ Grünbaum, *Geometry and Chronometry in Philosophical Perspective*, pp. 316–17.

senses. Thus, suppose that two light signals are *jointly* emitted from a point P on the periphery of a rotating disk and are each made to traverse that periphery in opposite directions so as to return to P. Then the two oppositely directed light pulses are indeed said to traverse *equal spatial distances*, i.e., the relation X is asserted for their paths on the strength of the obtaining of the relation RR. But the one light pulse which travels in the direction *opposite* to that of the disk's rotation with respect to frame I will travel a shorter path in I, and hence will return to the disk point P earlier than the light pulse traveling in the direction in which the disk rotates. Hence the round-trip times of light for these *equal* spatial paths will be *unequal*.⁵² (3) Whereas the spatial geometry yielded by a rigid rod on a stationary disk is Euclidean, such a rod yields a hyperbolic non-Euclidean geometry on a rotating disk.⁵³

What is the significance of these several results? We see that space intervals which are RR are called "congruent" ("X") in reference frames in which Y is relevant to their being *non-X*, no less than RR intervals are called "X" in those frames in which Y is relevant to their being X! Contrary to Giannoni's general contention, this important case from geometry and optics exhibits the pertinence of the distinction between semantic and non-semantic relevance. Here RR is positively

⁵² *Ibid.*, pp. 314–15.

⁵³ See, for example, C. Møller, *Theory of Relativity* (Oxford: Oxford University Press, 1952), Ch. VIII, §84.

For a rebuttal to John Earman's objections to ascribing a spatial geometry to the rotating disk, see the detailed account of the status of metrics in A. Grünbaum, "Space, Time and Falsifiability," Part I, *Philosophy of Science* 37 (Dec., 1970): 562–65 (chapter 16 of the present volume, pp. 542–545).

semantically relevant to being X. But Y has only *non*-semantical relevance to X, as shown by the fact that while Y is relevant to X in inertial systems, it is relevant to *non*-X in at least two kinds of non-inertial systems. And this refutes the generalization which Giannoni attributes to Duhem, viz., that there is parity of positive *semantic* relevance among properties such as RR and Y, so long as there are instances in which they are all merely positively *relevant* to X.

This lack of *semantic* parity between Y and RR with respect to "X" allows a semantically *stable* geometrical use of "X" on both the stationary and rotating disks, unencumbered by the *incompatibility* of the optical laws which relate the property Y to X in these two different reference frames. Indeed, by furnishing a *tertium comparationis*, this semantic stability confers *physical* interest on the contrast between the incompatible laws of optics in the two disks, and on the contrast between their two incompatible spatial geometries! For in *both* frames, RR is alike centrally semantically relevant to being X, i.e., in each of the two frames, rigid rod coincidences serve alike to determine what space intervals are assigned equal measures $ds = \sqrt{g_{ik} dx^i dx^k}$. This sameness of RR, which makes for the semantic stability of "X," need not, however, comprise a sameness of the correctional physical laws that enable us to allow computationally for the thermal and other deviations from rigidity.

I used certain formulations of relativity theory as a basis for the preceding account of the properties semantically and nonsemantically relevant to "X" ("spatially congruent"). But this account does not, of course, gain-say the legitimacy of alternative uses of "X" that would

issue in alternative, albeit physically equivalent, formulations. For example, the relativistic interpretation of the upshot of the Michelson-Morley experiment could alternatively have been stated as follows: In any inertial system, distances which are equal in the metric based on light-propagation are equal also in the metric which is based on rigid measuring rods.⁵⁴ I would say that this particular alternative formulation places Y on a par with RR as a *candidate* for being semantically relevant to "X." Moreover, I have strongly emphasized elsewhere that we are indeed free to formulate certain physical theories alternatively so as to make Y rather than RR semantically relevant to "X."⁵⁵

So much for matters pertaining to trivial fulfillments of D1.

ii. D2

It will be recalled that my example of a trivial fulfillment of D2 involved the use of an A' which is itself of the form $\sim H \vee O'$. In that example, the triviality did *not* arise from a violation of semantic stability. Instead, the fulfillment of D2 by the formal validity of the statement $[H \cdot (\sim H \vee O')] \rightarrow O'$ is only trivial because the entailment holds independently of the specific assertive content of H, unless a restrictive kind of entailment is used such as that of the system E of Anderson and Belnap. No matter what H happens to be about, and no matter whether H is substantively relevant to the specific content of O' or not, O' will be entailed

⁵⁴ This formulation is given in E. T. Whittaker, *From Euclid to Eddington* (London: Cambridge University Press, 1949), p. 63.

⁵⁵ Grünbaum, *Geometry and Chronometry . . .*, *passim*.

by H via the particular A' used here. Hence H does *not* serve to explain via A' the facts asserted by O', thereby satisfying D2 here only trivially, *if at all*.

As Philip Quinn has shown,⁵⁶ one of the fallacies which vitiates J. W. Swanson's purported syntactical proof that D2 holds interestingly⁵⁷ is the failure to take cognizance of this particular kind of semantic trivialization or even outright falsity of D2.

This brings us to the second of my previously published claims. The latter was that in its non-trivial form, D2 has not been demonstrated by means of D1. Having dealt with several *necessary* conditions for the non-triviality of A', I therefore went on to comment on a possible *sufficient* condition by writing:

I am unable to give a formal and completely general *sufficient* condition for the *nontriviality* of A'. And, so far as I know, neither the originator nor any of the advocates of the D-thesis [D2] have ever shown any awareness of the need to circumscribe the class of *nontrivial* revised auxiliary hypotheses A' so as to render the D-thesis [D2] interesting. I shall therefore assume that the proponents of the D-thesis intend it to stand or fall on the kind of A' which we would all recognize as *non-trivial in any given case*, a kind of A' which I shall symbolize by A'_{nt}.⁵⁸

And I added that D2 was *undemonstrated*, since it does *not* follow from D1 that

$$(\exists A'_{nt}) [(H \cdot A'_{nt}) \rightarrow O'].$$

⁵⁶ Philip Quinn, "The Status of the D-Thesis," Section III, *Philosophy of Science*, vol. 36, No. 4, 1969.

⁵⁷ J. W. Swanson, "On the D-Thesis," *Philosophy of Science*, vol. 34 (1967), pp. 59-68.

⁵⁸ Grünbaum, "The Falsifiability of a Component of a Theoretical System," p. 278.

Mary Hesse discusses my charge that D2 is a non-sequitur, as well as my erstwhile objections to D1, which I shall qualify below. And she addresses herself to a particular example of mine which was of the following kind: the *original* auxiliary hypothesis A ingredient in the conjunction $H \cdot A$ is highly confirmed quite separately from H, but the conjunction $H \cdot A$ does entail a presumably incorrect observational consequence. Apropos of this example, she writes:

. . . Grünbaum admits that . . . 'A is only more or less highly confirmed' (p. 289)—a significant retreat from the claim of truth. He thinks nevertheless that it is crucial that confirmation of A can be *separated* from that of H. But surely this is not sufficient for his purpose. For as long as it is not empirically demonstrable, but only likely, that A is true, it will always be possible to reject A in order to save H. . . . Which is where the Duhem thesis came in. It must be concluded that D [the D-thesis] still withstands Grünbaum's assaults. . . .⁵⁹

It is not clear how Hesse wants us to construe her phrase "in order to save H," when she tells us that "it will always be possible to reject A in order to save H." If she intends that phrase to convey merely the same as "with a view to attempting to save H," then her objection pertains only to my critique of D1, which is not now at issue. And in that case, she cannot claim to have shown that the D-thesis "still withstands Grünbaum's assaults" with respect to my charge of non-sequitur against D2. But if she interprets "in order to save H" in the sense of D2 as meaning the same as "with the assurance

⁵⁹ Hesse, *The British Journal for the Philosophy of Science* 18: 334–35.

that there exists an A'_{nt} that permits H to explain O' ," then I claim that she is fallaciously deducing D2 from the assumed truth of D1. I quite agree that if D1 is assumed to be true *and* IF THERE EXISTS AN A'_{nt} such that H does explain O' via A'_{nt} , then D1 would guarantee the feasibility of upholding A'_{nt} or H (but not necessarily both) as true. But this fact does not suffice at all to establish that there *exists* such an A'_{nt} ! Thus, my charge of non-sequitur against the purported deduction of D2 from D1 does *not* rest on the complaint that the Duhemian cannot assure *a priori* being able to marshal supporting evidence *for* A'_{nt} .

Thus, if Hesse's criticisms were intended, in part, to invalidate my charge of non-sequitur against D2, then I cannot see that they have been successful. And since Philip Quinn has shown⁶⁰ that D2, in turn, does *not* entail D1, I maintain that D1 and D2 are logically independent of one another.

This concludes the statement of my reasons for not retracting either of the first two of my three erstwhile claims, which were restated early in this section. Hence we are now ready to consider whether there are any known counterexamples to D1.

Thus, our question is: Is there any component hypothesis H of some theory T or other such that H can be sufficiently isolated so as to lend itself to separate observational falsification? It is understood here that H *itself* does not have any observational consequences. The example which I had adduced to answer this question affirmatively in earlier publications was drawn from physical geometry. I now wish to re-examine my

⁶⁰ University of Pittsburgh Doctoral Dissertation (1970).

claim that the example in question is a counterexample to D1. In order to do so, let us consider an arbitrary surface S . And suppose that the Duhemian wants to uphold the following component hypothesis H about S : if lengths $ds = \sqrt{g_{ik}} dx^i dx^k$ are assigned to space intervals by means of *rigid* unit rods, then Euclidean geometry is the metric geometry which prevails physically on the given surface S . Before investigating whether and how this hypothesis H might be falsified, we must be mindful of an important assumption which is ingredient in the antecedent of its if-clause.

That antecedent tells us that the numbers furnished by *rigid* unit rods are to be centrally semantically relevant to the interval measures ds of the theory. Thus, it is a necessary condition for the *consistent* use of rigid rods in assigning lengths to space intervals that any collection of two or more initially coinciding unit solid rods of whatever chemical constitution can thereafter be used *interchangeably* everywhere in the region S , *unless* they are subjected to independently designatable perturbing influences. Thus, the assumption is made here that there is a concordance in the coincidence behavior of solid rods such that no inconsistency would result from the subsequent interchangeable use of initially coinciding unit rods which remain *unperturbed* or "rigid" in the specified sense. In short, there is concordance among rigid rods such that all rigid rods *alike* yield the same metric ds and thereby the same geometry. Einstein stated the relevant empirical law of concordance as follows:

All practical geometry is based upon a principle which is accessible to experience, and which we will now try

to realise. We will call that which is enclosed between two boundaries, marked upon a practically-rigid body, a tract. We imagine two practically-rigid bodies, each with a tract marked out on it. These two tracts are said to be "equal to one another" if the boundaries of the one tract can be brought to coincide permanently with the boundaries of the other. We now assume that:

If two tracts are found to be equal once and anywhere, they are equal always and everywhere.

Not only the practical geometry of Euclid, but also its nearest generalisation, the practical geometry of Riemann and therewith the general theory of relativity, rest upon this assumption.⁶¹

I shall refer to the empirical assumption just formulated by Einstein as "Riemann's concordance assumption," or, briefly, as "R." Note that the assumption R predicates the concordance among rods of different chemical constitution on their being rigid or free from perturbing influences. Perturbing influences are exemplified by inhomogeneities of temperature and by the presence of electric, magnetic and gravitational fields. Thus, two unit rods of different chemical constitution which initially coincide when at the same standard temperature will generally experience a destruction of their initial coincidence if brought to a place at which the temperature is different, and the amount of their thermal elongation will depend on their chemical constitution. By the same token, a wooden rod will shrink or sag more in a gravitational field than a steel one. Since such perturbing forces produce effects of different magnitude on

⁶¹ A. Einstein, "Geometry and Experience," in *Readings in the Philosophy of Science*, H. Feigl and M. Brodbeck, eds. (New York: Appleton-Century-Crofts, 1953), p. 192.

different kinds of solid rods, Reichenbach has called them "differential forces." The perturbing influences which qualify as differential are linked to designatable sources, and their presence is certifiable in ways other than the fact that they issue in the destruction of the initial coincidence of chemically different rods. But the set of perturbing influences is open-ended, since physics cannot be presumed to have discovered *all* such influences in nature by now. This open-endedness of the class of perturbing forces must be borne in mind if one asserts that the measuring rods in a certain region of space are free from perturbing influences or "rigid."

It is clear that if our surface *S* is indeed free from perturbing influences, then it follows from the assumption *R* stated by Einstein that *any* two rods of different chemical constitution which initially coincide in *S* will coincide everywhere else in *S*, independently of their respective paths of transport. This result has an important bearing on whether it might be possible to falsify the putative Duhemian hypothesis *H* that the geometry *G* of our surface *S* is Euclidean.

For suppose now that *S* does satisfy the somewhat idealized condition of actually being macroscopically free from perturbing influences. Let us call the auxiliary assumption that *S* is thus free from perturbing influences "*A*." Then the conjunction *H* · *A* entails that measurements carried out on *S* should yield the findings required by Euclidean geometry. Among other things, this conjunction entails, for example, that the ratio of the periphery of a circle to its diameter should be π on *S*. But suppose that the surface *S* is actually a sphere rather than a Euclidean plane and that the measurements carried out on *S* with presumed rigid rods yield

various values significantly *less* than π for this ratio, depending on the size of the circle. How then is the Duhemian going to uphold his hypothesis H that if lengths ds on S are measured with rigid rods, the geometry G of S will be Euclidean so that all circles on S will exhibit the ratio π independently of their size?

Clearly, he will endeavor to do so by denying the initial auxiliary hypothesis A, and asserting instead the following: the rods in S are not rigid after all but are being subjected to perturbing influences constituted by differential forces. Obviously, just as in the case of Dicke's refutation of Einstein's theory, the falsification of H itself requires that the truth of A be established, so that Duhem would not be able to deny A. Suppose that we seek to establish A on the strength of the following further experimental findings: However different in chemical appearance, all rods which are found to coincide initially in S preserve their coincidences everywhere in S to within experimental accuracy independently of their paths of transport. How well does this finding establish the truth of A, i.e., that S is *free* from all perturbing influences which would act differentially on rods of diverse chemical constitution?

Let us use the letter "C" to denote the statement that whatever their chemical constitution, any and all rods invariably preserve their initial coincidences under transport in S . Then we can say that Riemann's concordance assumption R states the following: If A is true of *any* region S , then C is true of that region S . Clearly, therefore, C follows from the conjunction $A \cdot R$, but *not* from A alone. Hence if the observed preservation of coincidence or concordance in S establishes anything here, what it does establish at best is the conjunction

A · R rather than A alone, since I would *not* wish to argue against the Duhemian holist that any finding C which serves inductively to establish a conjunction A · R must be held to establish likewise each of the conjuncts separately.⁶² This fact does not, however, give the Duhemian a basis for an objection to my attempt to use the observed concordance in order to establish A. To see this, recall that R must be assumed by the Duhemian if his claim that S is Euclidean is to have the physical significance which is asserted by his hypothesis H. Since the Duhemian wants to uphold H, he could not contest R but only A. Hence, in challenging my attempt to establish A, the Duhemian will *not* be able to object that the observed concordance of the rods in S could establish only the conjunction A · R rather than establishing A *itself*. The issue is therefore one of establishing A. And for the reasons given in Einstein's statement of R above, R is assumed both by the Duhemian, who claims that our surface S is a Euclidean plane, and by the anti-Duhemian, who maintains that S is a sphere.

The Duhemian can argue that the observed concordance cannot conclusively establish A for the following two reasons: (1) He can question whether C can establish A, even if the *apparent* concordance is taken to establish the *universal* statement C indubitably, and (2)

⁶² By the same token, if *alternatively* the initial coincidence of the rods had been observed to *cease conspicuously* in the course of transport, the latter observations could *not* be held to falsify A in isolation from R. This inconclusiveness with respect to the falsity of A itself would obtain in the face of the *apparent* discordances, even if the latter are taken as *indubitably* falsifying the claim C of universal physical concordance. For C is entailed by A · R rather than by A alone.

he can challenge the initial inference from the seeming concordance to the physical truth of C. Let me first articulate each of these two Duhemian objections in turn. Thereafter, I shall comment critically on them.

Objection 1

The Duhemian grants that the mere presence of a *single* perturbing influence of significant magnitude would have produced differential effects on chemically different rods. Thus, he admits that the presence of a *single* such influence would be incompatible with C. But he goes on to point out that C is compatible with the joint presence of several perturbing influences of possibly unknown physical origin which just *happen* to act as follows: different effects produced on the rods by each one of these perturbing influences are of just the right magnitude to combine into the same *total* deformation of each of the rods. And all being deformed *alike* by the collective action of differential forces *emanating from physical sources*, the rods behave in accord with C. In short, the Duhemian claims that such a *superposition* of differential effects is compatible with C and that it therefore cannot be ruled out by C. Hence he concludes that C cannot be held to establish the truth of A beyond question. Since this Duhemian objection is based on the possibility of conjecturing the specified kind of superposition, I shall refer to it as "the superposition objection."⁶³

⁶³ This kind of objection is raised by L. Sklar, "The Falsifiability of Geometric Theories," *The Journal of Philosophy* 64 (1967): 247-53. Swanson, "On the D-Thesis," presents a model universe for which he conjectures superposition of differential ef-

Objection 2

This second Duhemian objection calls attention to the fact that in inferring C from the *apparent* concordance, we have made the following several inductive inferences: (i) We have taken differences in chemical *appearance* to betoken actual differences in chemical constitution, (ii) we have taken apparent coincidence to betoken actual coincidence, at least to within a certain accuracy, and (iii) we have inductively inferred the *invariable* preservation of coincidence by *all* kinds of rods everywhere in S from only a limited number of trials. The first two of these three inferences invoke collateral hypotheses. For they rest on presumed laws of psychophysics and neurophysiology to the following effect: certain appearances, which are contents of the awareness of the human organism, are lawfully correlated with the objective physical presence of the respective states of affairs which they are held to betoken. Hence let us say that the first two of these three inferences infer a conclusion of the form "Physical item P *has* the property Q" from the premise "P *appears* to be Q," or that they infer *being* Q from *appearing* to be Q.

These objections prompt two corresponding sets of comments.

fects *in name only*. For, as Philip Quinn has noted (in "The Status of the D-Thesis"), Swanson effectively adopts the convention that even in the *absence* of perturbing influences, all initially coinciding rods will be held to be *non-rigid* by being assigned lengths that *vary alike* with their positions and/or orientations. Thus, Swanson renders the superposition conjecture physically empty by ignoring the crucial requirement that the alleged differential effects are held to emanate individually from physical sources. Swanson's model universe will be discussed further in footnote 72.

1. It is certainly true that A is *not* entailed by the conjunction $R \cdot C$. For R is the conditional statement "If A , then C ." Since A does not follow deductively from $C \cdot R$, the premise $C \cdot R$ *cannot* be held to rule out the rival superposition conjecture DEDUCTIVELY. Thus, given R , C *deductively allows* the superposition conjecture as a *conceivable* alternative to A . Nevertheless, we shall now see that in the context of R , the inductive confirmation of A by C is so enormously high that A can be re-regarded as *well-nigh established* by $R \cdot C$. Since my impending inductive argument will make use of Bayes' theorem, it will be unconvincing to those who question the applicability of that theorem to the probability of a *hypothesis*. For example, Imre Lakatos registered an objection in this vein.⁶⁴

I shall adopt the non-standard probability notation employed by H. Reichenbach⁶⁵ and W. C. Salmon⁶⁶ to let " $P(I, M)$ " mean "the probability *from* I to M " or "the probability of M , given I ." In the usual notations, the two arguments I and M are reversed.

Recalling that R says "If A , then C " (for any region S), we wish to evaluate the probability $P(R \cdot C, A)$ of A , given that $R \cdot C$. To do so, we use a special form of Bayes' theorem⁶⁷ and write

$$P(R \cdot C, A) = \frac{P(R, A)}{P(R, C)} \times P(R \cdot A, C).$$

⁶⁴ See Imre Lakatos and Alan Musgrave, eds., *Criticism and Growth of Knowledge* (New York: Cambridge University Press, 1970), p. 187. But if he countenances entertaining the superposition conjecture $\sim A$ on the grounds that A cannot be said to be highly confirmed, how will this $\sim A$ escape being a *metaphysical* proposition?
See Appendix §49

⁶⁵ H. Reichenbach, *The Theory of Probability* (Berkeley: University of California Press, 1949).

⁶⁶ W. C. Salmon, *The Foundations of Scientific Inference* (Pittsburgh: University of Pittsburgh Press, 1967), p. 58.

C follows *deductively* from $R \cdot A$. Hence C is guaranteed by $R \cdot A$, so that

$$P(R \cdot A, C) = 1.$$

Thus, our formula shows that our desired posterior probability $P(R \cdot C, A)$ will be very high, if the ratio $P(R, A) / P(R, C)$ is itself close to 1, as indeed it will now turn out to be. Needless to say, the *ratio* of these probabilities can be close to 1, even though they are individually quite small.

We can compare the two probabilities in this ratio with respect to magnitude without knowing their individual magnitudes. To effect this comparison, we first consider the conditional probability of C, if we are given that R is true while A is *false*. We can now see that this probability $P(R \cdot \sim A, C)$ is very low, whereas we recall that $P(R \cdot A, C) = 1$. For suppose that A is false, i.e., that S is subject to differential forces. Then C can hold only in the *very rare* event that there *happens* to be just the right superposition. Incidentally in the case of our surface S, there is no evidence at all for the existence of the *physical sources* to which the Duhemian needs to attribute his superposed differential effects. Hence we are confronted here with a situation in which the Duhemian wants C to hold under conditions in which superposed differential forces are actually operative, although there is no independent evidence whatever for them.⁶⁸

⁶⁷ For this form of the theorem, see Reichenbach, *Theory of Probability*, p. 91, equation (6).

⁶⁸ This lack of evidence is here construed as obtaining at the given time. An alternative construal in which there avowedly would *never* be any such evidence at all is tantamount to an admission that A is true after all, or that $\sim A$ is physically empty. See Philip Quinn ["The Status of the D-Thesis,"] for a discussion of the import of this point for L. Sklar's *particular* superposition objection.

Since $P(R \cdot \sim A, C)$ is very low, we can say that the probability $P(R, C)$ of concordance among the chemically different rods is *only slightly greater, if at all*, than the probability $P(R, A)$ of the occurrence of a region S which is free from differential forces. And this comparison among these probabilities holds here, notwithstanding the fact that there may be as yet undiscovered kinds of differential forces in nature, as I stressed above. It follows that $P(R \cdot C, A)$ is close to 1. And since

$$P(R \cdot C, \sim A) = 1 - P(R \cdot C, A),$$

the conditional probability of $\sim A$ is close to zero.⁶⁹

Thus, the Duhemian's first objection does not detract, and indeed may not be intended to detract from the fact that A is *well-nigh* established by $R \cdot C$.

⁶⁹ H. Putnam claims to have shown in "An Examination of Grünbaum's Philosophy of Geometry," in *Philosophy of Science*, The Delaware Seminar, vol. 2, B. Baumrin, ed., (New York: Interscience Publishers, 1963), pp. 247-55, that a superposition of so-called gravitational, electromagnetic, and interactional differential forces can always be invoked here to assert $\sim A$. In particular, he is committed to the following: Suppose A is assumed for a region S (say, on the basis of C) and that a particular metric tensor g_{ik} then results from measurements conducted in S with rods that are presumed to be rigid in virtue of A . Then, says Putnam, we can *always* assert $\sim A$ of S by reference to the specified three kinds of forces. And we can do so in such a way that rods corrected for the effects of these forces would yield any desired different metric tensor g'_{ik} . The latter tensor would enable the Duhemian to assert his chosen geometric H of S .

But I have demonstrated elsewhere in detail [Grünbaum, *Geometry and Chronometry in Philosophical Perspective*, Ch. III, §9] that the so-called gravitational, electromagnetic and interactional forces which Putnam wishes to invoke in order to assert $\sim A$ do *not* qualify *singly* as differential. Hence Putnam's proposed scheme cannot help the Duhemian.

2. The second objection noted that being Q is inferred inductively from appearing to be Q in the case of chemical difference and of coincidence. Furthermore, it called attention to the inductive inference of the *universal* statement C from only a limited number of test cases. What is the force of these remarks?

Let me point out that the claim of mere chemical *difference* does *not* require the definite identification of the particular chemical constitution of each of the rods as, say, iron or wood.⁷⁰ Now suppose that the apparent chemically relevant differences among most of the tested rods are striking. And note that they can be further enriched by bringing additional rods to S . Would it then be helpful to the Duhemian to postulate that all of these *prima facie* differences in chemical appearance mask an underlying chemical identity? If he wishes to avoid resorting to the fantastic superposition conjecture, the Duhemian might be tempted to postulate that the great differences in chemical appearance belie a true crypto-identity of chemical constitution. For by postulating his crypto-identity, he could hope to render the observed concordance among the rods unproblematic for $\sim A$ *without* the superposition conjecture. Crypto-identity might make superposition dispensable because the Duhemian might then be able to assert that S is subject to only *one* kind of differential force.⁷¹ But

⁷⁰ For some pertinent details, see Grünbaum, "The Falsifiability of a Component of a Theoretical System," pp. 286–87.

⁷¹ It is not obvious that the Duhemian could make this assertion in this context with impunity: even in the presence of only one kind of differential force, there may be rods of one and the same chemical constitution which exhibit the hysteresis behavior of ferromagnetic materials in the sense that their coincidences at a given place will depend on their paths of transport. Cf. Grünbaum, *Geometry and Chronometry* . . . , p. 359.

there is no reason to think that the crypto-identity conjecture is any more probable than the incredible superposition conjecture. And either conjecture may try a working scientist's patience with philosophers.⁷²

As for the inductive inferences of physical coincidence from apparent coincidence, the degree of accuracy to which it can be certified inductively is, of course, limited and macroscopic. But what would it avail the Duhemian to assume in lieu of superposition the imperceptibly slight differential deformations which are compatible with the limited accuracy of the observations of coincidence? Surely the imprecision involved here does not provide adequate scope to modify A sufficiently to be able to reconcile the substantial observed deviations

⁷² Swanson, "On the D-Thesis," pp. 64–65, maintains that when the Duhemian contests A in my geometric example, he would *not* invoke a crypto-identity conjecture, as I have him do, but rather a conjecture of a crypto-difference among *prima facie* chemically identical rods. And that crypto-difference might arise from a kind of undiscovered isomerism. He says:

It is not the assertion of the chemical *difference* of two rods that the D-theorist would claim to be theory-laden. Grünbaum is certainly right in showing that the D-theorist could not argue for *that*. Rather, it is the assertion of their *identity* that would be theory-laden. For suppose that according to a given chemical theory T_1 , two rods a and b were found to be chemically identical to one another. . . . We can now construe T_1 as some sort of pre-"isomeric" chemistry . . . and then go on to imagine a post-"isomeric" T_2 such that upon the fulfillment of certain tests . . . it is found that $a \neq b$ But if identity of a and b is thus theory-laden, it does little good to argue that non-identical rods can be clearly discriminated.

Swanson then considers a model universe of three rods, two of which are *prima facie* chemically identical though crypto-different, while being pairwise objectively and perceptibly different from the third. And he then introduces two (ghostlike) perturbing influences whose *superposition* imparts the same total deformation to each of the three rods. Thus, in the context of his

of the circular ratios from π with the Euclidean H that he needs to save.

What of the inductive inference of the universal statement C from only a limited number of test cases? Note that C need not be established in its full universality in order to create inductive difficulties for the superposition conjecture and to provide strong support for A. For as long as all the chemically different rods actually tested satisfy C, the Duhemian must make his superposition conjecture or crypto-identity assumption credible with respect to *these* rods. Otherwise, he could not reconcile the measurements of circular ratios furnished by them on S with his Euclidean H.

superposition model, Swanson attaches significance to the fact that the first two of his three rods, a_1 and a_2 , are crypto-different.

But nothing in his argument from superposition depends on the hypothesis that a_1 and a_2 are crypto-different instead of identical, nor does his superposition model gain anything from that added hypothesis. Surely the differential character of his two perturbing influences would not be gainsaid by the mere superposition that a_1 and a_2 are chemically identical and that either of the two perturbing influences therefore affects them alike. If there is superposition issuing in the same total deformation, as contrasted with the absence of differential perturbations, then chemically identical rods can exhibit the same total deformation no less than conspicuously different or crypto-different ones! Far from lending added plausibility to the superposition conjecture, Swanson's introduction of the hypothesis of crypto-difference merely compounds the inductive felonies of the two.

In fact, Swanson overlooked the non-sequitur which I did commit apropos of crypto-identity in my 1966 paper. As is clear from the present essay, the Duhemian could couple his denial of A with the superposition conjecture *rather than* with the hypothesis of crypto-identity. It was therefore incorrect on my part to assert in 1966 that, when denying A, the Duhemian "must" assert crypto-identity (See "The Falsifiability of a Component of a Theoretical System," p. 287, *item* (3)).

The Duhemian might point out that I have inferred A inductively in a two-step inductive inference, which first infers at least instances of the physical claim C from apparent coincidences of *prima facie* different kinds of rods, and then proceeds to infer A inductively from these instances of C, coupled with R. And he might object that the relation of inductive support is not transitive. It is indeed the case that the relation of inductive support fails to be transitive.⁷³ But my inductive inference of A does not require an intermediate step via C, for cases of *apparent* coincidence, no less than cases of *physical* coincidence, are correlated with $A \cdot R$, and R is no more in question here than before. Thus, given R, the inference can proceed to A directly via the assumption that most cases of apparent coincidence of *prima facie* chemically different rods are correlates of cases of A.⁷⁴ And my inductive inference of A is thus predicated on this correlation.

In conclusion, we can now try to answer two questions: (i) To what extent has our verification of A been separate from the assumption of H?, and (ii) to what extent has H *itself* been falsified? When I speak here of a hypothesis as having been "falsified" in asking the latter question, I mean that the *presumption* of its falsity

⁷³ For a lucid explanation of the *non-transitivity* (as distinct from *intransitivity*) of the relation of inductive support and of its ramifications, see W. C. Salmon, "Consistency, Transitivity, and Inductive Support," *Ratio* 7 (1965): 164-68.

⁷⁴ This claim of direct inductive inferability is made here only with respect to the specifics of this example. And it is certainly not intended to deny the existence of other situations, such as those discussed by R. C. Jeffrey, *The Logic of Decision* (New York: McGraw-Hill, 1965), pp. 155-56, in which a perceptual experience can reasonably prompt us to entertain a variety of relevant propositions.

has been established, *not* that its falsity has been established with certainty or irrevocably. This construal of "falsify" is, of course, implicit in the presupposition of my question that there are different degrees of falsification or differences in the extent to which a hypothesis can be falsified.

(i) Our verification of A did proceed in the context of the assumption of R. And while we saw that R is ingredient in the Euclidean H of our example, as specified, R is similarly ingredient in the rival hypothesis that our surface S is not a Euclidean plane but a spherical surface. Thus, our verification of A was separate from the assumption of the *distinctive* physical content of the particular Duhemian H.

(ii) Duhem attributed the inconclusiveness of the falsification of a component hypothesis H to the legitimacy of denying instead any of the collateral hypotheses A which enter into any test of H. Our analysis has shown that *the denial of A is legitimate precisely to the extent that its VERIFICATION suffers from inductive uncertainty*. Moreover, in each of our examples of attempted falsification, the inconclusiveness is attributable *entirely* to the inductive uncertainty besetting the following two *verifications*: the verification of A, and the verification of the so-called observation statement which entails the *falsity* of the conjunction $H \cdot A$. In short, the inconclusiveness of the falsification of a component H derives wholly from the inconclusiveness of verification. And the falsification of H itself is inconclusive or revocable in the sense that the falsity of H is *not* a *deductive* consequence of premises *all* of which can be known to be true with certainty.

Hence if the falsification of H denied by Duhem's D1

is construed as *irrevocable*, then I agree with Mary Hesse⁷⁵ that my geometrical example does not qualify as a counterexample to D1. But I continue to claim that it does so qualify, if one requires only the *very strong presumption* of falsity. And to the extent that my geometrical example does falsify the A'_{nt} which D2 invokes in conjunction with the H of that example, the example also refutes D2's claim that A'_{nt} can justifiably be held to be true.

Subject to an important *caveat* to be issued presently, I maintain, therefore, that there are cases in which we can establish a strong presumption of the falsity of a component hypothesis, although we cannot falsify H in these cases beyond any and all possibility of subsequent rehabilitation. Thus, I emphatically do allow for the following possibility, though not likelihood, in the case of an H which has been falsified in my merely presumptive sense: A daring and innovative scientist who continues to entertain H, albeit as part of a new research program which he envisions as capable of vindicating H, *may* succeed in incorporating H in a theory so subsequently fruitful and well-confirmed as to retroactively alter our assessment of the initial falsification of H. And my *caveat* is that my conception of the falsification of H as establishing the strong presumption of falsity is certainly *not* tantamount to a stultifying injunction to any and all imaginative scientists to cease entertaining H forthwith, whenever such falsification obtains at a given time! Nor is this conception of falsification to be construed as being committed to the historically false assertion that no inductively *unwarranted* and daring

⁷⁵ Hesse, *The British Journal for the Philosophy of Science* 18:334.

continued espousal of H has ever been crowned with success in the form of subsequent vindication.

APPENDIX

One might ask: Are there no *other* cases at all in which H can be justifiably rejected as *irrevocably* falsified by observation? On the basis of a schema given by Philip Quinn, it may *seem* that if we can ignore such uncertainty as attaches to our observations when testing H, then this question can be answered affirmatively for reasons to be stated presently. When asserting the *prima facie* existence of this kind of irrevocable falsification as a fact of logic, Quinn recognized that its relevance to actual concrete cases in the pursuit of empirical science is at best very limited. Specifically, Quinn invites consideration of the following kind of logical situation. Let the observation statement O_3 be pairwise *incompatible* with each of the two observation statements O_1 and O_2 , so that $O_3 \rightarrow \sim O_1$, and $O_3 \rightarrow \sim O_2$. And suppose further that

$$[(H \cdot A) \rightarrow O_1] \cdot [(H \cdot \sim A) \rightarrow O_2].$$

Assume also that observations made to test H are taken to establish the truth of O_3 . Then we can deduce via *modus tollens* that

$$\sim (H \cdot A) \cdot \sim (H \cdot \sim A).$$

Using the law of excluded middle to assert $A \vee \sim A$, we can write

$$(A \vee \sim A) \cdot [\sim (H \cdot A) \cdot \sim (H \cdot \sim A)].$$

The application of the distributive law and the omission

of one of the two conjuncts in the brackets from each of the resulting disjuncts yields

$$[A \cdot \sim (H \cdot A)] \vee [\sim A \cdot \sim (H \cdot \sim A)].$$

Using the principle of the complex constructive dilemma, we can deduce $\sim H \vee \sim H$ and hence $\sim H$.

Although the assumed truth of the observation statement O_3 allows us to deduce the falsity of H in this kind of case, its relevance to empirical science is at best very limited for the following reasons.

In most, if not all, cases of actual science, the conjunction $H \cdot \sim A$ will not be rich enough to yield an observational consequence O_2 , if the conjunction $H \cdot A$ does yield an observational consequence O_1 . The reason is that the mere denial of A is not likely to be sufficiently specific in content. What, for example, is the observational import of conjoining the *denial* of Darwin's theory of evolution to a hypothesis H concerning the age of the earth? Thus this feature at least severely limits the relevance of this second group of logically possible cases to actual science.

Furthermore, the certainty of the conclusion that H is false in this kind of case rests on certainty that the observational statement O_3 is true. Such certainty is open to question. But *if* we can ignore our doubts on this score, then it would seem that H can be held to have been falsified beyond any possibility of subsequent rehabilitation. Yet, alas, John Winnie has pointed out that despite appearances to the contrary, the H in Quinn's schema does not qualify as a component hypothesis after all.

For Winnie has noted that Quinn's basic premise

$[(H \cdot A) \rightarrow O_1] \cdot [(H \cdot \sim A) \rightarrow O_2]$ permits the deduction of

$$H \rightarrow (O_1 \vee O_2).$$

Thus, if $O_1 \vee O_2$ can be held to qualify as an observationally testable statement, then we find that H entails it *without* the aid of the auxiliary A or of $\sim A$. But in that case, Winnie notes that H does not qualify as a constituent of a wider theory in the sense specified in our statement of D1 above.⁷⁶

⁷⁶ I am indebted to Philip Quinn and Laurens Laudan for reading a draft of this paper and suggesting some improvements in it. I also wish to thank Wesley Salmon and Allen Janis for some helpful comments and references.

The rapid production of new results in physics requires me to supplement the doubts stated on page 580, concerning the solar oblateness postulated by Dicke. In January, 1971, after this paper had been written, it was pointed out by A. P. Ingersoll and E. A. Spiegel ("Temperature Variation and the Solar Oblateness," *The Astrophysical Journal* 163 (1971), pp. 375 and 381) that Dicke and Goldenberg's solar oblateness measurement may be explained alternatively by an equatorial temperature excess of 30°K, located in the high photosphere and low chromosphere, and that presently available data do not permit to distinguish such temperature variation from true oblateness. Furthermore, in the same issue of *The Astrophysical Journal* (January 15, 1971, p. 361), B. Durney has noted that pole-equator differences in temperature, in turn, appear naturally as a result of the interaction of convection with rotation.

See Append.
§ 50

CAN AN INFINITUDE OF OPERATIONS BE PERFORMED IN A FINITE TIME?

In his book *Philosophy of Mathematics and Natural Science*, the renowned mathematical physicist Hermann Weyl raised the following question: Is it *kinematically* feasible that a machine carry out an *infinite* sequence of distinct operations in a finite time? And he gave a conditional answer to it as follows: If a machine obeying the principles of classical kinematics cannot carry out a denumerable infinity of operations in a finite time, then the received interpretation of the classical mathematical theory of motion is beset by one of Zeno's kinematical paradoxes. Thus, Weyl refers to Achilles' task of traversing a unit space interval by successively traversing the infinite series of decreasing subintervals of lengths

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots, \frac{1}{2^n} \dots (n = 1, 2, 3, \dots).$$

And Weyl writes (p. 42):

'The remark that the successive partial sums . . . of the series . . . do not increase beyond all bounds but converge to 1, by which one nowadays thinks to dispose of the paradox, is certainly relevant and elucidating. Yet, if the segment of length 1 really consists of infinitely many sub-segments of lengths $1/2, 1/4, 1/8, \dots$, as of 'chopped-off' wholes, then it is incompatible with the character of the infinite as the 'incompletable' that Achilles should have been able to traverse them all. If one admits this possibility, then there is no reason why a machine should not be capable of completing an infinite sequence of distinct acts of decision within a finite amount of time; say, by supplying the first result after $1/2$ minute, the second after another $1/4$ minute, the third $1/8$ minute later than the second, etc. In this way it would be possible, provided the receptive power of the brain would function similarly, to achieve a traversal of all natural numbers and thereby a sure yes-or-no decision regarding any existential question about natural numbers!

Weyl's contention has generated a literature alleging the kinematic absurdity of hypothetical machines which are presumed to be capable of carrying out an infinite sequence of operations *in a finite time*. Hereafter I shall often speak of $\aleph_0$ operations instead of saying 'an infinite

This chapter was presented as a Monday Lecture at the University of Chicago in April 1968. It appears here by permission of the University of Chicago.

sequence of operations'. These so-called 'infinity machines' have included a device that would print *all* of the digits of an infinite decimal on a finite strip of paper, a machine that would recite the names of all the natural numbers in a suitable code, a lamp that would be switched on and off $\aleph_0$ times, and a device that would transfer a marble back and forth $\aleph_0$ times.

Charges of paradox have been levelled against infinity machines by such various authors as G. J. Whitrow, M. Black, C. S. Chihara, and J. F. Thomson.¹ Their allegations of absurdity pertain to the distinctively kinematical aspects of the $\aleph_0$ operations which are to be executed. Hence the physiological, chemical, electrical or other feasibility of the $\aleph_0$ operations will not be at issue, unless it has a bearing on the assessment of their kinematic possibility. But if there are physiological, electrical or other *non*-kinematic grounds for deeming the infinity machines to be physically or technically impossible, is it then not an idle exercise to inquire whether they are *kinematically* paradoxical? I think not. Of course, I grant instantly that such an inquiry may not be productive for engineering. But I hope to show that it is indeed illuminating with respect to the kinematical component of the class of physical theories which assert the mathematical continuity of space and time, as relativity theory does, for example. Even standard quantum theory employs continuous space and time variables, although it has, of course, repudiated the well-defined particle trajectories of Newtonian and relativistic mechanics.

In this paper, I shall offer an *affirmative* though *multiply qualified* answer to Weyl's question concerning the performability of $\aleph_0$ operations in a finite time. But I shall endeavour to show that Weyl's own mildly conditional affirmative answer is too strong, viz. his claim that if the unit interval traversed by Achilles in unit time 'really consists' of $\aleph_0$ geometrically decreasing subintervals, then a machine must be able to complete an infinite sequence of 'distinct acts of decision' (e.g. calculations).

It will turn out that the fundamental kinematic issues posed by Weyl's contention can be confronted by comparing the motion of Achilles, who runs *continuously* at an average unit velocity, with the motion of another runner who is presumed to traverse the same unit space interval in the same unit time but runs *intermittently* as follows: he interrupts his motion by $\aleph_0$ pauses of rest whose successive durations have the following geometrically decreasing magnitudes:

$$\frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \dots,$$

and so on *ad infinitum*. Thus, the latter intermittent motion, which I shall

¹ For details, see my (1968) *Modern Science and Zeno's Paradoxes*, ch. II, § 4, and my (1968a, p. 396).

call the '*staccato*' run, will serve as our prototype of the $\aleph_0$ operations which are to be executed by the infinity machines in a finite time.

Preparatory to inquiring whether the *staccato* run is beset by kinematic absurdities, I must explain in detail my reasons for denying Zeno's charge of paradox against the mathematical description of Achilles' *uninterrupted* motion through a unit space interval. I shall refer to Achilles' smooth run as the '*legato*' motion in order to distinguish it from the intermittent motion of his *staccato* running mate. And for our entirely non-historical purposes, I take Zeno's challenge to be the following: Kinematic theory tells us convincingly enough that a *legato* runner moving with average unit velocity can traverse a unit space interval in unit time. But, this theory *also* asserts that Achilles' *legato* run involves, among other things, an *infinite* sequence of submotions through a progression of non-overlapping spatial subintervals of respective lengths $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, . . . , and so on. It is true that the successive *durations* of these submotions can be such as to form a series which suitably converges to zero. Nonetheless, Zeno continues, the *legato* motion cannot be completed in a finite time, since the theory tells us that the elapsing of a unit time interval involves the successive elapsing of an *endless* progression of subintervals. And he contends that kinematics thereby contradicts itself, because its *unending* sequences of submotions render the *completion* of any motion *temporally unintelligible*.

THE LEGATO RUN

In honour of Zeno, let us apply the name 'Z-sequence' to an infinite progression of intervals of space or time whose successive magnitudes are $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, . . . , and so on. For the sake of arithmetic simplicity, I shall follow Zeno's procedure and assume that the successive *durations* of the *legato* runner's submotions form a Z-sequence just as the subintervals of space which are covered by these submotions. Thus, it is being assumed that the *legato* runner's *average* velocity is unity in each of the $\aleph_0$ subintervals traversed by him up to the terminal instant of his motion. But we shall see later in our discussion of the *staccato* run that this *constancy* of the average velocity makes for a kinematically dubious discontinuity in the velocity function at the terminal instant. Fortunately, the *legato* runner can alternatively be assumed to run such that his velocity function yields average velocities in the subintervals that converge to zero at the terminal instant. But let us disregard this arithmetically more complicated *legato* motion until we discuss the *staccato* run. And let us furthermore adopt the terminology used by Vlastos (1966, p. 95) and refer to the traversal of any of the

spatial *subintervals* of our Z-sequence as 'making a Z-run'. Vlastos notes that as commonly used, the term 'run' individuates uniquely the physical action to which it applies, much as 'heart-beat' does. And he points out that in *this* sense of 'run', the runner's *legato* traversal of the Z-sequence could only be described as a *single* run and not as having involved a denumerable infinity of 'Z-runs'. But clearly, in order to traverse the unit interval in one smooth and uninterrupted 'run' in the ordinary sense, the runner must—among other things—traverse all the members of the Z-sequence and, in the latter sense, make $\aleph_0$ Z-'runs'. To distinguish between these two quite different uses of the noun 'run', Vlastos writes 'run_a' for the single motion which we can perceive with our unaided senses in daily life contexts, and 'run_b' for the kind relevant to the Z-sequence of kinematics.

Human awareness of time exhibits a positive threshold or minimum. This fact can now be seen to have a consequence of fundamental relevance to the appraisal of Zeno's argument. For it entails that *none* of the infinitely many temporal subintervals in the progression whose magnitude is less than the human *minimum perceptibilium* can be *individually experienced as elapsing* in a way that does *metrical* justice to its actual lesser duration. To succeed, the attempted individual *contemplation* of all the subintervals would require a denumerable infinity of mental acts, each of which requires or exceeds a positive minimum duration. *Thus, we do not experience these subintervals individually as elapsing in a metrically faithful way. Instead, we gain our metrical impression of duration in this context from the time needed by our mental acts of contemplation and not from the respective duration numbers which we associate intellectually with the contemplated subintervals when performing these mental acts.* And the resulting compelling feeling that an infinite time is actually needed to accomplish the traversal in turn insinuates the deducibility of this paradoxical result from the theory of motion. Specifically, the existence of a duration-threshold of time awareness guarantees that there is a *positive lower bound* on the duration of any run_a. And this fact enters into several of the following fallacies here committed by Zeno in his so-called Dichotomy paradox:

(1) Zeno's claim that the progression of Z-runs_b requires an infinite future time is made plausible by a tacit appeal to our awareness that $\aleph_0$ runs_a would indeed last forever, because there is a positive lower bound on the duration of any run_a. The threshold governing our acts of awareness likewise induces the feeling that after the first instant of the motion, a unique next event must happen in the motion *and* that there must be a unique next-to-the-last event that happens before the final instant of the

motion, if there is to be a final instant at all. But since the successive sub-intervals converge to zero by decreasing geometrically, the threshold-governed one-by-one contemplation which Zeno invites cannot be metrically faithful to the actual physical durations of the contemplated sub-intervals. And our intuitive time awareness rightly boggles at *experiencing* each of $\aleph_0$ subintervals of time as elapsing individually. But, justified though it is, Zeno *illicitly* trades on this boggling. For it cannot detract from the following crucial fact: *any and every* one of the $\aleph_0$ temporal subintervals of the motion is over by the end of one unit of time. To see this, note that for *every* n , the sum of the first n terms of the geometric series of duration numbers

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots, \frac{1}{2^n}$$

is given by a total duration less than 1, namely by the sum

$$S_n = 1 - \left(\frac{1}{2}\right)^n \quad (n = 1, 2, 3, \dots).$$

And note further that here I have invoked the undisputed fact that the clock measures of time intervals are *finitely additive* to claim that one unit of time is the least upper bound on the total duration of all the $\aleph_0$ subintervals of the motion. It follows that both distributively and *collectively* all $\aleph_0$ temporal subintervals of the motion elapse within one unit of time. The justification for this conclusion becomes further apparent upon becoming cognisant of the next error, by which Zeno buttresses his conclusion that the runner would *never* reach his destination.

(2) With respect to the relation of temporal precedence, the set comprising the temporal subintervals of the progression *and* the instant of the runner's arrival at his destination has the form of an *infinite* progression followed by a last element. And this ordered set is said to be of ordinal type $\omega + 1$. Furthermore, the instant of the runner's arrival at his destination point i does *not* belong to *any* of the temporal subintervals of the progression. Thus, the closed time interval required by the runner's completed motion consists of all the instants belonging to any of the subintervals of the progression *and* of the instant of arrival at the point i . By failing to include the instant of arrival, the membership of the subintervals of the progression fails to exhaust the entire membership of the closed time interval required by the completed motion. Indeed, the union of the subintervals of the progression constitutes a half-open time interval precisely because it fails to include the terminal instant of the motion.

Zeno illicitly exploits the fact that it is *logically* impossible for the *terminal* instant of the motion to belong to any of the subintervals of the unending progression, since it is *later* than all of them. Zeno appeals to

this fact to infer wrongly that *there cannot be* any terminal instant at which the runner reaches his destination. Once we have become victimised by our threshold of time awareness, we are set up to commit this otherwise transparent fallacy. In this way, Zeno seeks to lend further credence to his claim that the union of the subintervals of the progression is of infinite duration. But what can be deduced from the logical impossibility of finding the terminal instant in any of the subintervals forming the unending progression? The failure of the terminal instant to belong to the union of the $\aleph_0$ subintervals amounts to no more than that this instant is not to be found in the half-open time interval which has been left half-open by *excluding the terminal instant*; the half-openness of the resulting time interval does *not* show that the union of the temporal subintervals must be of infinite *duration* just because that union has no terminal instant, and just because the infinite progression of subintervals has no last member. For the terminal instant is the *earliest* instant *following every* instant belonging to *any* subinterval of the unending progression, while the durations of these subintervals suitably converge to zero. The non-existence in the progression of a last subinterval during which the motion would be completed does not preclude the existence of an instant later than all the subintervals which is the last instant of the motion.

In the case of Zeno's arithmetically simple example, this state of affairs expresses itself *arithmetically* in the following compound way:

(1) If the runner departs at $t = 0$, then corresponding to the non-existence of a last temporal subinterval of the motion in the progression, the respective times by which he has traversed the successive subintervals of the Z-sequence are given by the *infinite* sequence

$$\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \frac{31}{32}, \dots, \frac{2^n - 1}{2^n}, \dots \quad (n = 1, 2, 3, \dots).$$

(2) Although the number 1 is not a member of this infinite sequence of time numbers, the arithmetic limit of this infinite sequence on the number axis is constituted by the number 1, which is the time coordinate of the last instant of the motion and represents the total duration of the union of the subintervals belonging to the progression.

(3) The runner traverses ever shorter subintervals of the unit race course in proportionately ever shorter subintervals of time, thereby travelling at constant average speed.

What then are we to think of the charge that the arithmetic theory of limits has been lifted uncritically out of the context of its legitimate application to physical *space* and adduced irrelevantly in an effort to refute Zeno's allegations of *temporal* paradox? We saw that the mathe-

mathematical apparatus of the theory of limits is ordinally and metrically no less appropriate to physical time than it is to physical space. Note that I have *not* invoked the *arithmetic* theory of limits *as such* to dismiss the allegation that kinematical theory entails *temporal* paradoxes. Instead, my contention has been that we are *justified* by the ordinal and metrical structure of physical time in applying that arithmetical theory *and* that Zeno's specific deductions of metrical contradictions in the Dichotomy paradox are each vitiated by fallacies which I am engaged in pointing out.¹

The highly misleading role played by Zeno's one-by-one contemplation of the members of his progression becomes conspicuous upon noting the following fact: *It would even take us forever to contemplate one-by-one the progression of DURATIONLESS instants which divide one temporal subinterval from the next, and yet the durational measure of this progression of instants is ZERO within standard physical theory!* By the same token, the fact that our contemplation of the $\aleph_0$ subintervals would last forever is not a basis for concluding that the union of the progression of them would be of infinite duration. In summary, Zeno would have us infer that the runner can *never* reach his destination, just because (1) in a finite time, we could not possibly contemplate one by one *all* the subintervals of the progression, and (2) for purely *logical* reasons, we could not possibly find the terminal instant of the motion in *any* of the $\aleph_0$ subintervals of the progression, since the terminal instant is not a member of any of them. But it is altogether fallacious to infer Zeno's conclusion of infinite duration from these two premises.

The recent literature on Zeno continues to provide illustrations of the intellectual havoc resulting from an irrelevant *though tacit* appeal to the fact that there is a positive lower bound on the duration of any single mental act of ours such as *conscious counting*. Thus, G. J. Whitrow seems to have engaged in precisely such an unwitting appeal in his endeavour to show that the mathematical continuity which we attribute to finite intervals of space cannot similarly be attributed to physical time without thereby generating logical antinomies. After stating that 'We must not assume that . . . in *time*, any infinite sequence of operations can be performed' Whitrow (1961, p. 148) considers the consequences of assuming that our *legato* runner passes through the entire progression of positions envisaged by Zeno as the respective termini of his subintervals. Whitrow

¹ In the wake of Zeno, A. N. Whitehead, W. James and others have claimed that the elements of time must succeed one another *discretely* and hence cannot intelligibly possess the denseness property of the linear mathematical continuum, which requires that between *any* two instants, there exists at least one other. I must refer to my (1968) chap. II, § 2, B and C for a statement of my reasons for rejecting this contention.

invites us to assume that in so doing, the runner would *number* all these positions consecutively and concludes that then the runner's task would involve exhausting 'the infinite set of positive integers by *counting*'.¹ Thus Whitrow conjures up the image of conscious counting in a manner akin to Zeno's illicit appeal to the eternity of one-by-one contemplation in the Dichotomy. But it is wrong to identify the metrical features of the process of conscious counting (say, in English) with those of traversing Zeno's progression of points in a finite interval. For we must not allow our human psycho-neurophysiology of time awareness to intrude itself and to victimise us in this mathematically most misleading way.

We recall that Zeno's *legato* runner requires $\frac{1}{2}$ unit of time to traverse the first $\frac{1}{2}$ of the unit space interval, $\frac{1}{4}$ unit of time to traverse the next $\frac{1}{4}$ of the distance, $\frac{1}{8}$ unit of time for the next $\frac{1}{8}$ of the distance, and so on. Let us now examine the different motion of the *staccato* runner, who departs simultaneously and runs on an equal and parallel race track.

Our description of the *staccato* run so far did not specify the magnitudes of the successive distances traversed by the *staccato* runner during the times allotted to his intermittent runs. For the sake of arithmetic simplicity, I shall first treat the *staccato* run by taking the successive distances to be proportional to the times of the intermittent runs by a factor of 2. Thus all $\aleph_0$ intermittent motions of the *staccato* runner are first assumed to proceed at the *same* average velocity. But our assessment of the kinematic feasibility of the *staccato* run will *not* be made to rest on the arithmetically simple case of constant average velocity, any more than in the case of the *legato* run. For we shall see that this simple case involves the following kinematically problematic feature: at the terminal instant of the runner's arrival at his final destination, his velocity exhibits a *finite* discontinuity. Hence after having treated this arithmetically simple case, I shall be concerned to note that alternatively, the *intermittent* runs of the *staccato* runner no less than the uninterrupted sequence of *legato* runs could demonstrably proceed at *suitably decreasing* average velocities so as *not* to exhibit any kind of discontinuity. In this way, the *staccato* run will turn out to be just as unproblematic kinematically as the *legato* run.

Having entered this caution, I shall first confine my attention entirely to the arithmetically simple case.

¹ In here using everyday words such as *doing* and *things* in technical contexts, we are fully alerted against such confusions as misidentifying a *run_b* as a *run_a*, no less than when we use technical terms such as *work* and *energy* in physics. We need to use language to describe the physical process constituting the *legato* runner's traversal of the total interval. And in determining whether this process can occur in a finite time, as described, we need to heed the commitments of ordinary language only to the extent of guarding against being victimised or stultified by them.

THE STACCATO RUN

By contrast to the *legato* runner, his *staccato* mate is required to traverse each subinterval of the Z-sequence in half the time needed by the *legato* runner and then to wait for the latter to catch up with him before traversing the next Z-interval. Specifically, the *staccato* runner takes $\frac{1}{4}$ of a unit of time to traverse the first Z-interval of length $\frac{1}{2}$ and rests for an equal amount of time; then he takes $\frac{1}{8}$ of a unit of time to traverse the second Z-interval of length $\frac{1}{4}$ and rests for an equal amount of time, and so on *ad infinitum*. Thus, the *staccato* runner interrupts his motion by $\aleph_0$ pauses of rest whose successive durations are

$$\frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \dots$$

And he suspends his motion at each of his Z-stops just long enough to enable his *legato* colleague Achilles to catch up with him. It might *appear* quite legitimate to require the *staccato* runner to plant a flag at each of his Z-stops while he pauses at them and awaits his friend Achilles. Indeed, even the *legato* runner Achilles has been called upon in the literature to engage in flag-planting at each of the progression of points terminating the subintervals of the Z-sequence on his track. This demand has been addressed to the *legato* runner on Zeno's behalf in the recent literature, notwithstanding the fact that Achilles spends only a mathematical instant of zero duration at each of the specified space points while running uninterrupted.

But if the *staccato* runner is to carry out the $\aleph_0$ operations of starting and stopping during a finite time, kinematic theory will not allow him to carry out the $\aleph_0$ operations of planting a flag at each of the designated Z-stops. To see this, note first that the erection of a flag at each of the $\aleph_0$ Z-stops would presumably require the *staccato* runner to translate his own limbs and rotate the flag *each* time into the vertical direction through a minimum positive distance, however small. And in that case, the *staccato* runner would have to perform $\aleph_0$ upward (or downward) *equal* minimal spatial displacements during ever shorter times at boundlessly increasing average velocities. Thus, he would have to effect a *spatially infinite* total displacement of his own limbs and of the flags during a finite time in the following manner: The successive *maximum* vertical velocities of his limbs required to erect the flags consecutively in ever shorter times would *increase boundlessly* with time up to the instant $t = 1$ at which *he comes to rest* at his destination, since the successive average velocities would have to increase boundlessly. But such a motion has two kinematically objectionable features: (a) at the instant $t = 1$ of arrival at the destination point P, the motion of his hands violates the requirement that the

THE STACCATO RUN

By contrast to the *legato* runner, his *staccato* mate is required to traverse each subinterval of the Z-sequence in half the time needed by the *legato* runner and then to wait for the latter to catch up with him before traversing the next Z-interval. Specifically, the *staccato* runner takes $\frac{1}{4}$ of a unit of time to traverse the first Z-interval of length $\frac{1}{2}$ and rests for an equal amount of time; then he takes $\frac{1}{8}$ of a unit of time to traverse the second Z-interval of length $\frac{1}{4}$ and rests for an equal amount of time, and so on *ad infinitum*. Thus, the *staccato* runner interrupts his motion by $\aleph_0$ pauses of rest whose successive durations are

$$\frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \dots$$

And he suspends his motion at each of his Z-stops just long enough to enable his *legato* colleague Achilles to catch up with him. It might *appear* quite legitimate to require the *staccato* runner to plant a flag at each of his Z-stops while he pauses at them and awaits his friend Achilles. Indeed, even the *legato* runner Achilles has been called upon in the literature to engage in flag-planting at each of the progression of points terminating the subintervals of the Z-sequence on his track. This demand has been addressed to the *legato* runner on Zeno's behalf in the recent literature, notwithstanding the fact that Achilles spends only a mathematical instant of zero duration at each of the specified space points while running uninterrupted.

But if the *staccato* runner is to carry out the $\aleph_0$ operations of starting and stopping during a finite time, kinematic theory will not allow him to carry out the $\aleph_0$ operations of planting a flag at each of the designated Z-stops. To see this, note first that the erection of a flag at each of the $\aleph_0$ Z-stops would presumably require the *staccato* runner to translate his own limbs and rotate the flag *each* time into the vertical direction through a minimum positive distance, however small. And in that case, the *staccato* runner would have to perform $\aleph_0$ upward (or downward) *equal* minimal spatial displacements during ever shorter times at boundlessly increasing average velocities. Thus, he would have to effect a *spatially infinite* total displacement of his own limbs and of the flags during a finite time in the following manner: The successive *maximum* vertical velocities of his limbs required to erect the flags consecutively in ever shorter times would *increase boundlessly* with time up to the instant $t = 1$ at which *he comes to rest* at his destination, since the successive average velocities would have to increase boundlessly. But such a motion has two kinematically objectionable features: (a) at the instant $t = 1$ of arrival at the destination point P, the motion of his hands violates the requirement that the

position of a body be a continuous function of the time, since the vertical position of a point on either hand does not approach any limit as $t \rightarrow 1$, and (b) at the terminal instant $t = 1$, the fluctuating *velocity* function of his hands has an instant of *infinite discontinuity*, since that velocity function is unbounded in every earlier neighbourhood of the terminal instant.

By contrast, if he is *not* required to plant any flags, then at $t = 1$, the (vertical and horizontal) *position* of a point on the *staccato* runner's hands or feet is a continuous function of time, and his horizontal velocity will fluctuate only between 0 and some fixed finite maximum value so as to yield an *average* 2 units for his intermittent horizontal motions. But since this *average* horizontal velocity maintains the *same* positive value up to the terminal instant of rest, the successive *peak* velocities cannot fall below this positive value and *cannot* converge to zero. Hence in this arithmetically simple example, the continuity of the horizontal position at the terminal instant $t = 1$ obtains alongside the following discontinuity: the horizontal velocity function exhibits an instant of *finite* discontinuity at $t = 1$, just as the graph of a step function has points of finite discontinuity. Moreover, since there is a minimum velocity change (from zero to the average) during each of the N_0 *ever shorter* time intervals, the horizontal *accelerations* increase (and decrease) boundlessly as $t \rightarrow 1$, and the acceleration function has an instant of infinite discontinuity at the terminal instant of rest-and-zero-acceleration.

Thus, we see that the requirement to plant a flag N_0 times at the Z-stops calls for a kinematically forbidden discontinuity in the vertical position along with an infinite discontinuity in the vertical velocity. On the other hand, the execution of the N_0 start-and-stop operations of the *staccato* run without flag-planting requires no discontinuity in the position. But if we demand that all the *staccato* motions proceed at the same average velocity, then this run involves a finite discontinuity in the horizontal velocity along with an infinity discontinuity in the horizontal acceleration.

It is clear, therefore, that the flag-planting is ruled out kinematically. But in view of the remaining discontinuities in the horizontal velocity and acceleration, our arithmetically simple kind of *staccato* and *legato* runs may well be kinematically problematic. As far as I know, books on classical or pre-quantum mechanics do not spell out whether motions involving these particular discontinuities are kinematically possible or not.

If the specified discontinuities in the horizontal velocity and acceleration are *not* impermissible, then it can now be shown to follow that the arithmetically simple *staccato* run is no less feasible kinematically than the corresponding *legato* motion. And in that case, the arithmetically simple *staccato* run can be consummated simultaneously with the *legato* run.

To see this, note that in the arithmetically simple case, the *staccato* runner at no time lags behind his *legato* colleague during the closed unit interval but is either ahead of him or abreast of him. While running within each of the Z -intervals, the *staccato* runner's average velocity is twice that of his *legato* colleague, but his overall average velocity for the total interval is equal to his colleague's velocity and is less than the velocity of light *in vacuo*. And if the aforementioned *finite* discontinuity in the horizontal velocity of the runners at $t = 1$ is not kinematically impermissible, then we can conclude the following: if the *legato* runner reaches his destination in 1 unit of time after traversing the Z -sequence, then so also does the *staccato* runner.

There *may*, of course, be specifically *dynamical* as distinct from kinematical—difficulties in effecting the infinitude of horizontal accelerations and decelerations required by the *staccato* runner's alternate starting and stopping. Indeed, calculation shows (cf. my 1968a, p. 401) that the total energy (work) expended by that runner in imparting the same average velocity to his body N_0 times is *infinite*. Thus, the runner would have required an infinite store of energy when he set out on his run. For he sustains N_0 *uncompensated* losses of kinetic energy in the decelerations, and the total magnitude of these losses is infinite. The availability of an infinite amount of energy is a matter of the *de facto* dynamical boundary conditions of the world, or at least, not a matter of kinematical law.

Fortunately, as Richard Friedberg has pointed out,¹ the intermittent runs no less than the uninterrupted sequence of *legato* runs can each proceed at suitably decreasing average velocities such that each runner's velocity, acceleration (and *all* of the higher time-derivatives as well) vary continuously throughout the closed unit time interval during which he traverses a unit distance. Thus, Friedberg's version of the *legato* and *staccato* runs obviates all of the kinematically and dynamically problematic features of the arithmetically simple example. In particular, the successive *peak* velocities and accelerations attained by Friedberg's *staccato* runner during the decreasing subintervals converge to zero as we approach the terminal instant.²

¹ Private communication from Professor Richard Friedberg of the Department of Physics at Columbia University.

² More specifically, let $K = \sum_{n=0}^{\infty} e^{-n^2}$, which equals about 1.38. Then Richard Friedberg's

N_0 intermittent runs cover the successive distances $\frac{e^{-n^2}}{K}$ ($n = 1, 2, 3, \dots$). He lets

and

$$g(x) = e^{-\csc^2 \pi x} \text{ for } 0 < x < 1$$

$$g(x) = 0 \quad \text{for } x \geq 1 \text{ and } x \leq 0.$$

And he notes that the function g and all of its derivatives are continuous for all real values of x .

[footnote continued on p. 641]

In the case of Friedberg's *legato* and *staccato* runs, we can therefore conclude without qualification that the *staccato* run is no less feasible kinematically than the *legato* motion and that both are indeed kinematically possible.

This conclusion has the following important consequences: Given that the pauses separating the individual traversals carried out by the *staccato* runner form a geometric progression whose terms converge to zero, it is *immaterial* to the traversability of the total unit interval in a finite time that the process of traversal consists of $\aleph_0$ motions separated by pauses of rest (as in the *staccato* run) instead of being one uninterrupted motion which *can be analysed* into an infinite number of submotions (as in the *legato* run). And if we wish to call the *staccato* runner's execution of the $\aleph_0$ *separate* motions 'doing infinitely many things,' then his performance shows that infinitely many things can be done in a finite time. Of course, if the pauses between the individual traversals of the *staccato* run were all *equal*, then this run could not be carried out in a finite time, no matter how small each of the equal pauses might be.

In view of the thresholds which govern the physiological reaction times

Upon putting

$$I \equiv \int_0^1 g(t) dt = \int_0^1 e^{-\csc^2 \pi t} dt,$$

he lets the *legato* runner have a velocity

$$v_l = \frac{g(t)}{I}.$$

It is then obvious that the distance $\int_0^1 v_l dt$ covered by the *legato* runner during the unit time interval is *one* unit.

Friedberg requires the *staccato* runner to run at the velocity

$$v_s = \frac{1}{IK} \sum_{n=0}^{\infty} 2^{n+2} \cdot g(2^{n+2}[t-1+2^{-n}])e^{-n^2}$$

which has the following properties: (a) *all* of the infinitely many terms of this sum vanish during all those time intervals when the *staccato* runner is required to rest, and (b) during the time of the n th intermittent run (beginning with $n = 0$), the positive velocity function v_s is given entirely by the n th term, since all the other terms of the sum from $n = 0$ to $n = \infty$ vanish for that time interval.

The minimum value 1 of $\csc^2 \pi x$ is yielded by $x = \frac{1}{2}$, so that the maximum value of $g(x)$ is given by $g(\frac{1}{2})$, which is $1/e$. Since the maximum value $g(\frac{1}{2})$ corresponds to the *temporal* midpoint of the n th intermittent run, the peak value P_n of v_s during the n th intermittent run occurs at its temporal midpoint. Hence Friedberg concludes that in the case of the velocity function v_s , the successive peaks in the intermittent runs (beginning with $n = 0$) are $P_n = 2^{n+2}/KeIe^{n^2}$ and converge to zero. And he shows that the distance $\int_0^1 v_s dt$ tra-

versed by the *staccato* runner during the unit time interval is likewise 1. The *staccato* runner moves and rests intermittently for the required geometrically decreasing times. But he gets ahead of the *legato* runner, who first catches up with him at $t = 1$, though Friedberg notes that the use of a more complicated formula would allow the *legato* runner to catch up as well at the earlier times $\frac{1}{2}, \frac{3}{4}, \dots$, as in the simplest example.

of the *staccato* runner and his times of *conscious* execution of a set of instructions, it is clear that this runner cannot be 'programmed' to perform the *staccato* run in accord with the required metrical specifications, when the times during which he is to run or rest become small enough to fall below his thresholds. But this fact does not vitiate my contention that, in principle, kinematically the *staccato* run as described is physically possible. For kinematic theory allows us to assume that a body's separate motions have the prescribed metrical properties.

In claiming to have shown that the *staccato* run is kinematically feasible, I have, of course, *not* offered a formal consistency proof of kinematic theory as supplemented by the description of the *staccato* run. It seems to me that instead, we can reasonably be content with the following kind of 'proof' that the *staccato* run is kinematically possible: A demonstration that the alleged deducibility of the paradoxical *infinite* duration is unfounded and that, given the kinematical principles of the theory along with the boundary conditions, the theory entails the *finitude* of the total duration of the *staccato* run.

We have been mindful of *ruling out* any flag planting or other marking processes that would require *any* discontinuous change in any of the three position coordinates of the *staccato* runner's limbs. But the *staccato* runner can be held to have 'marked' each one of a progression of space points by the act of stopping at each one for the prescribed length of time. If I may presume that this waiting at the stops qualifies as 'marking' them, then the *staccato* runner's total motion constitutes an important counter-example to one of the theses recently put forward by C. S. Chihara as part of his critical response to Weyl's comparison of the Z-run with the performance of an infinity machine. Chihara believes that for *logical* reasons the difference between Achilles' mere *legato* traversal of the interval and a runner's marking all the end points of the subintervals in the course of his journey makes for the difference between completeness in a finite time and requiring an infinite time. He says:

... to give a more intuitive characterization of the difference between Achilles' journey and Achilles' task of marking the end points, in the former case we start with the task and analyze it into an infinite sequence of stages, whereas in the latter case we start with the stages and define the task as that of completing the infinite sequence of stages. To complete the journey, one must simply perform a task which can be analyzed ad infinitum, but to complete the task of marking all the end points, one must really do an infinite number of things (Chihara (1965) p. 86).

But as we saw, the *staccato* runner does 'really do an infinite number of things' in what is kinematically a demonstrably *finite* time. It would

appear that here Chihara has misdiagnosed the source of the difference between completability in a finite time and requiring an infinite time. The *staccato* run is *not* one uninterrupted motion which can be merely *analysed* into a particular infinite set of submotions, as in the case of the *legato* run. Instead it consists of $\aleph_0$ motions separated by pauses of rest. And yet kinematically it is physically possible to complete it in a finite time in the context of classical physics.

We had to rule out the successive planting of $\aleph_0$ flags by the *staccato* runner as kinematically impossible, because it would involve a discontinuity in the position function. It turns out that in the case of all the other infinity machines mentioned at the outset, we must likewise predicate the assertion of their kinematic possibility on their at least *not* involving an illicit discontinuity in the position function. Let me merely illustrate the need for such a restriction by reference to the hypothetical machine that would print *all* of the digits of an infinite decimal such as the infinite decimal representation of the real number π .

If this machine is to achieve the printing of all the digits of π in a finite time, then the heights from which the press descends to the paper to print the successive digits may *not* be equal but must form a suitably decreasing series converging to zero. By means of this restriction, we can assure that the spatial magnitude of the successive tasks does *not* remain the same while the time available for performing them decreases toward zero. In this way, the kinematic demands made on the π -machine become more like those made on the *staccato* runner. My reason for requiring the heights of descent to converge to zero in a suitable fashion becomes apparent upon recalling the analysis I gave of the flag planting in the case of the *staccato* runner. If the heights of descent did not suitably converge to zero, the successive velocities of the press required for the printing would soon exceed the velocity of light in contravention of the special theory of relativity and would vary with time in a manner that is kinematically objectionable even in the context of the Newtonian theory. And the position function of the printing press would be discontinuous at the terminal instant. Here, no less than in the case of the *staccato* runner, I ignore the *dynamical* problems of programming the π -machine so that the successive spatially and temporally shorter descents of the press can be *triggered* as required. And I also disregard here whether an infinite time might not be required for the more complicated process by which the progression of digits might first have been *computed seriatim*, not to speak of the process of inserting into the printing press all the characters which are to print the digits.

Incidentally, I require that the widths of the successive numerals to be

printed converge to zero in such a way that *all* the $\aleph_0$ digits can be printed in a horizontal line on a *finite* strip of paper. In laying down this requirement, I blithely ignore as kinematically irrelevant the blurring of the digits on the paper through smudging of the ink when their widths become sufficiently small, not to speak of the need for ink droplets of width dimensions below those of an electron! Thus, this *particular* example of the printing press does ignore the atomic constitution of matter.

Given my requirement concerning the widths of the successive digits, the spatial array of the $\aleph_0$ digits no more requires an infinite space than the *unending* progression of Z-intervals which collectively fit into the space of a finite unit interval: As long as the sequence 3.1415926535 . . . is printed so that the successive widths of the digits converge to zero in the manner of Z-intervals, the question 'What does this array *look like at the right end?*' receives the same kind of answer as the corresponding question about the progression of Z-intervals. We must not make the misguided attempt to form a *visual* picture of the *open end* of a finite, half-open space interval, and we are aware that the metrically finite union of the Z-intervals is open at the right 'end' as is the total space interval formed by the progression of horizontally shrinking digits. Although we cannot visually picture the *non-existence* of a last digit or Z-interval, our very characterisation of the openness of the right end shows that we clearly understand in ordinal terms 'what that end looks like'. Just as the interval constituted by the union of the Z-intervals can be closed at the right end by the addition of a rightmost (last) point, so also, of course, can the interval formed by the horizontal cross section of the unending π -sequence.

In conclusion, let me emphasise that I gave reasons for *doubly* qualifying my assertion of the kinematic feasibility of $\aleph_0$ distinct operations in a finite time. First, I made my assertion contingent on the restriction that the spatial magnitude of the successive tasks *suitably* converge to zero. For example, a convergence to zero in the form of the divergent harmonic series $\frac{1}{n}$ ($n = 1, 2, 3, 4, \dots$) would not do. For the successive velocities required to traverse the distances given by this divergent series would increase boundlessly, and thus the velocity would exhibit an infinite discontinuity at the terminal instant of the unit time interval. Moreover, the velocity and acceleration may exhibit other discontinuities which are kinematically problematic, unless the velocity function meets the very stringent continuity requirements satisfied by Friedberg's particular *staccato* run. Weyl did not show that the operation of a machine which *calculates rather than merely prints* seriatim the digits of π can, in principle, avoid these various discontinuities. Hence it would seem that the following

claim by Weyl was either too strong or at least unfounded: Achilles can traverse *all* the *Z*-intervals in the manner of the received interpretation of kinematical theory *only* if an infinite sequence of calculations can be completed in a finite time.

REFERENCES

- CHIHARA, C. S. (1965) On the possibility of completing an infinite process. *Phil. Rev.* **74**, 74-87.
- GRÜNBAUM, A. (1968) *Modern Science and Zeno's Paradoxes*. London: Allen and Unwin Limited. Chap. 2, § 4. Pp. 78-120.
- GRÜNBAUM, A. (1968a) Are 'infinity machines' paradoxical? *Science* **159**, 396-406.
- VLASTOS, G. (1966) Zeno's race course. *J. Hist. Phil.* **4**, 95-108.
- WHITROW, G. J. (1961) *The Natural Philosophy of Time*. London: Nelson. P. 148.

IS THE COARSE-GRAINED ENTROPY OF CLASSICAL STATISTICAL MECHANICS AN ANTHROPOMORPHISM?*

1. INTRODUCTION

In Chapter 8, I claimed that the coarse-grained classical entropy statistics of certain ensembles of branch systems contribute to the 'arrow' of time. And in Chapter 22, §4, we shall transpose this theme to a relativistic space-time. But it has been charged that the coarse-grained entropy of a physical system is an anthropomorphism, incapable of a role in physically undergirding time's arrow. Hence it behooves us to face this charge. In the present chapter, I shall argue that the entropy in question can be validly construed in scientific realist fashion instead of being an anthropomorphism.

To introduce the issue posed in the title of our inquiry, let me caution against a possible misconstrual of the classification 'anthropomorphism' as a pejorative characterization of an attribute or entity postulated by a physical theory. It is clear enough that those who charge a *particular* concept employed in such a theory with being an anthropomorphism are *not* concerned to convey thereby the following truism: *Qua being a theory*, any known theory whatever is not only propounded and devised by man (or by a humanoid computer), but is also unavoidably *fallible* because of the postulational (inductive) risks inherent in the very logic of its intellectual construction. Thus, it is surely *not* the latter truism that E. T. Jaynes was interested in espousing when he wrote: "... entropy is an anthropomorphic concept, not only in the well-known statistical sense that it measures the extent of human ignorance as to the microstate. *Even at the purely phenomenological level, entropy is an anthropomorphic concept.* For it is a property, not of the physical system, but of the particular experiments you or I choose to perform on it."¹ What may be overlooked, however, is that those who level the charge of anthropomorphism at a particular, specified ingredient of a physical theory ought *not* to be saddled with the following claim:

For any and every item in such a theory, an unambiguous identification can always be made as to whether or not the item is an anthropomorphism within the postulational framework of the given theory. Nor should such a dichotomous claim be imputed to those who *deny* a *specific* charge of anthropomorphism in physics! This is not to say that the notion of anthropomorphism has not been employed in an avowedly dichotomous fashion, as for example in L. Kronecker's Pythagorean dictum "God made the integers; all the rest is the work of man."²

When considering the respective logical constitutions of the phenomenological and statistical entropies ascribed to a physical system, Jaynes singled out the extent of the role of human decision in the former and of human ignorance in the latter as touchstones for characterizing, or perhaps even indicting, them as anthropomorphisms. Whereas Jaynes was concerned to spell out his view of the bearing of human choice or decision on the status of the *phenomenological* entropy, I shall not deal with the latter kind of entropy at all. Instead, I shall focus on the significance of human volition in the logical constitution of the statistical coarse-grained entropy of classical statistical mechanics by addressing myself to the following question: Is the import of the role of human choice and decision for the *physical* credentials of that statistical entropy such as to sustain the charge of anthropomorphism against it?

As I shall set forth below in detail, such a charge might be made plausible by reference to the following fact: the occurrence and direction of a temporal change of the entropy $S = k \log W$ assigned to a given physical system depends essentially on *our human choice* of the *size* of the finite equal cells or boxes into which we partition the 6-dimensional position-velocity phase space (' μ -space'). In the pre-quantum classical statistical mechanics, there is enormous scope for the exercise of such choice in coarse-graining. For classical theory viewed action as a magnitude susceptible of having any value, and Planck's quantum of action h is unavailable in that theory to provide a *physical* basis for carving up the phase space into boxes each of which has the volume h^3 .³ It is precisely because there is such relatively great latitude for human decision in coarse-graining the specified phase space of *pre-quantum* statistical mechanics that I wish to examine the soundness of the charge of anthropomorphism as leveled against the statistical entropy of that

classical theory. It will then be a corollary of our analysis to determine whether the physical significance of that entropy can be validly impugned by regarding it as a measure of the extent of human ignorance as to the underlying microstate.

But before we can deal with the issue of anthropomorphism, we must give some detailed consideration to the aforementioned dependence of the temporal change of the entropy on the human decision as to the partitioning of the phase space into cells.

2. ENTROPY CHANGE AND ARBITRARINESS OF THE PARTITIONING OF PHASE SPACE

One of the things we must examine is the effect of *alternative* choices of a partition on the *direction* in which the entropy changes with time as the physical system develops. Let us briefly recall the meanings of some of the fundamental terminology amid explaining the special notation I shall use to indicate for each of the various pertinent quantities to what particular partition they are being referred.

Consider one particular system of particles of given finite total energy and given finite 3-dimensional volume. Call that system ' α '. Let the portion of 6-dimensional phase space whose occupation by the representative points of these particles is compatible with the given energy and physical dimensions be divided into a finite number m of equal cells $c_1, c_2, \dots, c_m$. Also let the number of particles which respectively occupy these various particular cells at a given time be $n_1, n_2, \dots, n_m$. Such a set of numbers $n_i (i = 1, 2, \dots, m)$ specifies the *macrostate* of the system by stating for each of the cells how many particles occupy it at the time. Thus the macrostate is characterized by the numerical *distribution* of the total finite number n of particles among the various particular cells, irrespective of which individual particles these may be. Suppose, for example, cells c_1 and c_2 were occupied respectively by unequal numbers of particles x and y at time t_1 , but then were occupied by y particles and x particles respectively at time t_2 . Then, since $x \neq y$, the macrostates (distributions) I and II at these two times t_1 and t_2 would be *different*, even though we assume that the numbers of particles in each of the remaining cells $c_3, c_4, \dots, c_m$ had not changed at all. For the macrostate does depend on the identity of the *cells* among which differing numbers

of particles are distributed, although the *macrostate* does not depend on the identity of the *particles* which are thus distributed.

The identity of the particles themselves is, of course, relevant to the *microstate* (arrangement, complexion). The latter is specified by stating for each of the n particles in which one of the m equal 6-dimensional cells of the coarse-grained partition its representative point lies at a given time. By thus being specified only to within the finite cells of the chosen partition, the micro-state as such is *relative to the chosen partition*. By contrast, the *precise* micro-state of classical mechanics is given by the exact or punctal values of the six position-velocity coordinates and is therefore *independent* of the chosen partition. To distinguish the latter kind of microstate from the former, I shall use the qualifying adjectives 'precise' or 'punctal'. Although the unqualified term 'microstate' will always refer to the partition-dependent, *non-punctal* kind of microstate, I shall occasionally emphasize that this relative kind of microstate is intended by using the adjective 'non-punctal' or 'coarse-grained'.

For any one chosen partition, each kind of macrostate or distribution is clearly constituted by a certain class of microstates: Any occurrence of one of the members of this class of microstates constitutes an occurrence of the macrostate in question. The number W of different microstates belonging to the particular macrostate specified by the set of distribution numbers $n_1, n_2, \dots, n_m$ is given by $W = n! / (n_1! \cdot n_2! \cdot \dots \cdot n_m!)$. (It will be convenient in the sequel to denote the product of the factorials in the denominator of this expression by ' $\pi n_i!$ '.) And the entropy S assigned to any one macrostate to which W microstates belong is $S = k \log W$. Since increases or decreases of W are tantamount respectively to increases or decreases of S , we can conveniently work with W below rather than with S when calculating the directions of entropy changes. Indeed, it should be borne in mind that it will be permissible for our particular purposes to talk about W and S interchangeably.

Since multiplication in the product $\pi n_i!$ is commutative, our aforementioned two *different* macrostates I and II are each assigned the same entropy S . Indeed, *any* distribution of the n particles of the given system that gives rise to the same value of the product π will be assigned the same value of the entropy. Thus, one and the same entropy state S_0 of a given system can correspond to *different* distributions or macrostates *each* of which has W_0 microstates belonging to it, where the value

of W_0 is given by $S_0 = k \log W_0$. Hence the number of different microstates which can underlie or belong to one and the same entropy state S_0 is given *not* by W_0 but rather by the *product* of W_0 with the number d_0 of different distributions having the same entropy S_0 .⁴

Now let A and B be two different kinds of precise, punctal microstates in which the given physical system x may be at different times. Denote a particular partition of the corresponding phase space by the numeral 1. And let $1A$ and $1B$ be the respective coarse, *non-punctal* microstates of the system *with respect to partition 1*, when it is in the respective punctal microstates A and B . We shall generally be interested in cases in which the *non-punctal* microstates are different from one another in kind, no less than the punctal ones.

If we consider the characterization of the system in states A and B to within an *alternative* partition 2, then $2A$ and $2B$ are the respective *non-punctal* microstates of the system with respect to partition 2. Clearly, the one precise microstate A gives rise to the two different *non-punctal* microstates $1A$ and $2A$, since the latter are specified only to within the cells of their different partitions. And similarly for B , $1B$ and $2B$.

Returning to partition 1, we note that $1A$ and $1B$ each uniquely specify the distribution (macrostate) to which they belong with respect to partition 1. These distributions may or may not be the same, but in either case we shall denote them by ' $D1A$ ' and ' $D1B$ ' respectively. Furthermore, we shall use ' W_{D1A} ' to represent the number of microstates belonging to $D1A$. For important reasons that we shall discuss later on, with respect to *the particular partition to which it pertains*, the quantity W_{D1A} does *triple duty* as follows: (i) As just noted, it is the number of different microstates belonging to the given distribution (macrostate) $D1A$, (ii) it is the (unnormalized) probability of occurrence of that distribution in the time ensemble of the macrostates which the system attains relatively to partition 1, i.e., the so-called thermodynamic probability of the macrostate in question, and (iii) it is a measure of the degree of *homogeneity* or evenness, equalization, disorder and 'well shuffledness' of the given macrostate $D1A$ in the class of macrostates which the system attains relatively to partition 1. This *triple significance* of the quantity W will turn out to have an important bearing on our central issue of anthropomorphism later on. Using the above notational devices, we

shall say that S_{D1A} and S_{D1B} are the entropies of macrostates $D1A$ and $D1B$ respectively, i.e., the entropies of the system relative to partition 1 when it is in the respective precise microstates A and B . By the same token, the macrostates which A and B constitute relatively to partition 2 have entropies S_{D2A} and S_{D2B} respectively.

We shall now illustrate the following important sets of facts concerning the bearing of the chosen partition on the direction of the entropy changes of the given system α :

(i) There are partitions 1 and 2 of the corresponding phase space as well as precise microstates A and B such that if the system evolves from A to B ,

$$S_{D1A} = S_{D1B} \quad \text{but} \quad S_{D2A} \neq S_{D2B}.$$

i.e., although the system does *not* change its entropy relatively to partition 1 in the course of its transition from A to B , its entropy does change relatively to partition 2. Thus, the system's two macrostates $D1A$ and $D1B$ with respect to partition 1 are equiprobable, whereas its two macrostates $D2A$ and $D2B$ with respect to partition 2 have different probabilities. There is no inconsistency between these various probabilities.

(ii) There are partitions 3 and 4 as well as precise microstates E and F such that if the system evolves from E to F ,

$$S_{D3E} < S_{D3F} \quad \text{but} \quad S_{D4E} > S_{D4F}.$$

i.e., although the entropy of the system *increases* relatively to partition 3 in the course of its transition from E to F , it *decreases* relatively to partition 4. Since these two entropy changes have opposite signs, the *probability ranking* of $D3E$ and $D3F$ will be opposite to that of $D4E$ and $D4F$. There is no inconsistency between these various probabilities.

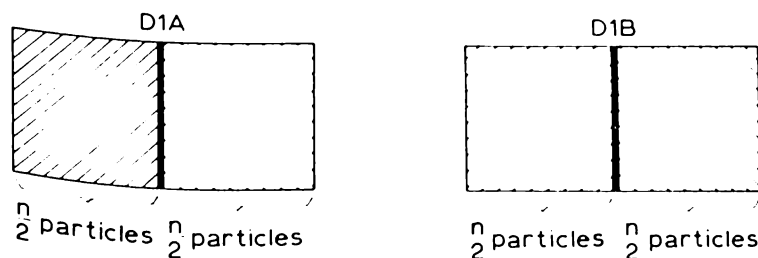
To make our diagramming and arithmetic quite simple, we shall represent the 6-dimensional phase space of the system α by a mere rectangle, and we shall choose four convenient numbers m of equal boxes (cells) to generate the four different partitions of the representative rectangle. Also, it will simplify our formulation without detriment to rigor to speak of the particles themselves rather than of their representative points as occupying certain cells in phase space.

Case (i)

We shall, of course, not label and depict the individual n particles, since we shall not diagram the precise microstates A and B but only their corresponding respective pairs of macrostates relative to partitions 1 and 2. The first partition will consist of two boxes ($m=2$), while the number m of boxes of partition 2 will be 4.

A macrostate relative to a given partition is specified *only to within the particular cells* of that partition. Hence the shading of a given box in the diagram will signify that the positive integral number of particles appropriate to the given distribution of particles is to be found *somewhere or other* in that box. If a given box is not shaded, then it contains no particles *anywhere* within it.

Let the precise microstate A be such that the first of the two boxes of partition 1 contains half of the particles *somewhere within it*, while the second, of course, contains the remaining half of the n particles *somewhere within it*. Thus $D1A$ is specified by the sequence $n/2, n/2$. Let B be a precise microstate different from A but such that the macrostate $D1B$ which it generates relatively to partition 1 is the same as $D1A$. Thus, we diagram:



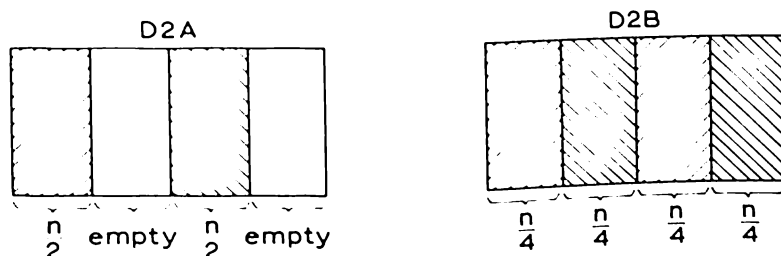
Since A and B generate identical distributions $D1A$ and $D1B$ relatively to partition 1, their thermodynamic probabilities relative to that partition will have the same value, viz., $n!/[(n/2)!]^2$. Hence relatively to partition 1, the system will have the same entropy when it is in punctal microstate A as when it is in the different punctal microstate B , and we can write

$$S_{D1A} = S_{D1B},$$

as claimed initially.

Turning to partition 2 ($m=4$), it is clear that A and B and the integral number n of particles can also be such as to generate the following quite

different distributions relatively to partition 2: In macrostate $D2A$, boxes 1, 2, 3 and 4 respectively contain $n/2$, 0, $n/2$ and 0 particles, whereas $D2B$ involves the uniform presence of $n/4$ particles in each of the four boxes. A and B can and must be so chosen that the requirement of energy conservation and other relevant constraints on the closed system α are indeed satisfied.



Hence we have

$$W_{D2A} = \frac{n!}{(n/2)! 0! (n/2)! 0!} = \frac{n!}{[(n/2)!]^2},$$

a value which, incidentally, is the same as that of W_{D1A} and W_{D1B} above. But

$$W_{D2B} = \frac{n!}{[(n/4)!]^4}.$$

Since $[(n/4)!]^4 < [(n/2)!]^2$, the quantities W pertaining to the different macrostates $D2A$ and $D2B$ satisfy the inequality $W_{D2A} < W_{D2B}$, so that we can write

$$S_{D2A} \neq S_{D2B},$$

as claimed initially.

We have illustrated the joint validity of the relations

$$S_{D1A} = S_{D1B} \quad \text{and} \quad S_{D2A} \neq S_{D2B}$$

by reference to a case in which $m=2$ for partition 1, while $m=4$ for partition 2. Thus, in our particular illustration the entropy of the system is *constant* for the transition from A to B relatively to the partition possessing the *smaller* number of cells. Thus, our particular illustration is one in

which the entropy *changes* relatively to the partition having the *larger* number of cells. But it is easy to provide another illustration of the joint validity of the above entropy relations in which the *constancy* of the entropy obtains relatively to that partition which has the *larger* number of cells. Hence consider the *general* case in which a transition of the system from *A* to *B* involves no entropy change relatively to one of two partitions while producing an entropy change relatively to the other. Then we can say that the *change* in entropy is *not* generally correlated with the partition containing the larger number of cells. In other words, the success or failure of one of two partitions in effecting an entropic differentiation between *A* and *B* is *not* generally a matter of the relative coarseness of the partition.

It is patent that there is no inconsistency between (1) the assertion $W_{D1A} = W_{D1B}$, when construed as claiming the *equiprobability* of the two macrostates *D1A* and *D1B*, and (2) the assertion $W_{D2A} \neq W_{D2B}$, when interpreted as affirming that the distributions *D2A* and *D2B* have *different* probabilities of occurrence in time. There are two different time ensembles of macrostates which our physical system *x* exhibits relatively to the two different partitions 1 and 2. In one of these two time ensembles, the two *different* distributions *D2A* and *D2B* respectively generated by *A* and *B* relatively to partition 2 can readily have *unequal* probabilities of occurrence, while the identical distributions *D1A* and *D1B* which *A* and *B* generate relatively to partition 1 will, of course, have one and the same probability. For the different probability *rankings*, i.e., the respective unequal and equal probabilities corresponding to these two partitions, are *not* ascribed to the pair of precise microstates *A* and *B* themselves but only to the two different pairs of macrostates *D1A*, *D1B* and *D2A*, *D2B* which are associated with the one pair *A* and *B* relatively to the two differing partitions. The consistency of the two different probability rankings corresponding to the two partitions we have diagrammed above is made palpable by these diagrams as follows: It would be a blatant physical falsehood to claim that the different macrostates *D2A* and *D2B* depicted in our second diagram are equiprobable, but it is plain that the two identical distributions of our first diagram are equiprobable.

A further important facet of the bearing of the chosen partition on the entropy changes of the system *x* will now emerge from illustrating the

oppositely directed entropy changes

$$S_{D3E} < S_{D3F} \quad \text{and} \quad S_{D4E} > S_{D4F}$$

stated as (ii) above.

Case (ii)

I am indebted to Allen Janis for the particular illustration which is about to follow and for a very helpful discussion of its significance for the entropy statistics of ensembles of systems.

Let partition 3 consist of 9 boxes, and let the precise microstates E and F as well as the number n of particles be such that the following two distributions are generated in the course of the system's transition from E to F : $D3E$ is constituted by the presence of *all* of the n particles in the first box of partition 3, while $D3F$ is a uniform distribution of the particles among the 9 boxes, so that each of the 9 boxes of partition 3 contains $n/9$ particles. Hence we have

$$W_{D3E} = \frac{n!}{n!} = 1 \quad \text{and} \quad W_{D3F} = \frac{n!}{[(n/9)!]^9}.$$

Without computing the value of the second of these two probabilities, it is evident that many more than 1 non-punctal microstate belong to $D3F$, while only 1 such microstate belongs to $D3E$. Therefore, $W_{D3E} < W_{D3F}$, and $S_{D3E} < S_{D3F}$, as claimed initially.

Without specifying the actual number m of cells in partition 4, let that number (as well as the states E and F) be such that $D4E$ and $D4F$ have the following respective properties: In macrostate $D4E$, each box of the comparatively fine partition 4 contains *at most one* particle, whereas in macrostate $D4F$, each cell of this partition contains either two particles or none. Note that in the case of $D4F$, the n particles are so distributed that there are $n/2$ cells each of which contains exactly two particles while the remaining cells are each empty. Hence we have

$$W_{D4E} = \frac{n!}{1} = n! \quad \text{and} \quad W_{D4F} = \frac{n!}{(2!)^{n/2}}.$$

Evidently $W_{D4E} > W_{D4F}$, so that $S_{D4E} > S_{D4F}$, as claimed initially.

Therefore, during the transition of the system α from E to F , the entropy relative to partition 3 *increases* whereas the entropy with respect

to partition 4 *decreases*. For reasons of the kind discussed under Case (i) above, the probability ranking $W_{D3E} < W_{D3I}$ is entirely *consistent* with the *opposite* probability ranking $W_{D4E} > W_{D4I}$.

Joseph Camp has given me an illustration of the existence of partitions 5 and 6, and of a temporal sequence of precise microstates *A*, *B* and *C* such that both

$$W_{D5A} < W_{D5B} < W_{D5C} \quad \text{and} \quad W_{D6B} < W_{D6C} < W_{D6A}.$$

We have set forth the *relativity to the chosen partition* of the very occurrence *and* direction of an entropy change in *one* particular physical system *z* during its transition from one punctal microstate to another. This relativity to the chosen partition for a single system *z* raises the important question whether the *statistics* of temporal entropy change in large but finite *ensembles* of like closed systems are similarly partition-dependent and might thereby be claimed to depend indirectly on human choices.

To deal with this question, let us be more precise as to the nature of the ensembles of systems to which it pertains. The systems in each and every ensemble will be alike: Each system will consist of, say, the same kind of gas confined to regions of three-dimensional space having the same volume, and possessing the same total energy. In *one* of these ensembles, say K_1 , to which our one system *z* above belongs, each system is in an initial state which is the same kind of distribution (macrostate) as *D3E* relatively to partition 3, and the same kind of distribution as *D4E* relatively to partition 4. But the precise microstates which underlie these same initial distributions in the various members of K_1 are emphatically *different* in kind from *E* in all except *z*, i.e., the precise microstates all differ from one another in kind. Moreover, the respective non-punctal microstates which generate the initial macrostate *D4E* in the various systems of K_1 will *each* be a *random sample* of the FINITE set of W_{D4E} different microstates compatible with *D4E* in the pertinent member of K_1 . I recognize that the concept of a random sample or of a random selection is not wholly unproblematic as characterized in the following standard kind of definition: "A sample obtained by a selection of items from the population is a *random sample* if each item in the population has an equal chance of being drawn. *Random* describes a method of drawing a sample, rather than some resulting property of the sample discoverable after the obser-

vance of the sample.”⁵ But I assume that such difficulties as may beset the concept of a random sample in its application to *finite* sets – which is the application I have made of it here – can be resolved as follows: The clarified concept of random sample accords with the physical correctness of the statistical claim that I shall make below concerning the entropic behavior of the ensembles of systems which we are engaged in discussing. In the case of the initial distribution D_{3E} , there is only one underlying microstate for each system since $W_{D_{3E}} = 1$, and hence this random sample requirement is trivially satisfied. All the members of K_1 have, of course, the same initial entropy $S_{D_{3E}}$ relative to partition 3, and also the same initial entropy $S_{D_{4E}}$ with respect to partition 4.

A *second* ensemble, K_2 , resembles K_1 in that the initial macrostate of its members relative to partition 3 has the same entropy $S_{D_{3E}}$, and similarly for the initial macrostate of entropy $S_{D_{4E}}$ relative to partition 4. But the members of K_2 differ from those of K_1 in that their initial macrostates relative to our two partitions are constituted by distributions which differ in kind from D_{3E} and D_{4E} . Except for the latter difference in the characteristics of K_1 and K_2 , the other statements I made about the initial conditions governing the members of K_1 apply, *mutatis mutandis*, to those of K_2 . By the same token, we can consider further ensembles of systems $K_3, K_4, \dots, K_n$, where n is a finite integer, in *each* of which the systems start out in entropy state $S_{D_{3E}}$ and in entropy state $S_{D_{4E}}$ but with the following difference between the various ensembles: Relative to a *given* partition, the kind of distribution which constitutes the initial entropy state of the systems in *one* of the ensembles *differs* from the kind of distribution constituting the self-same initial entropy state of the systems in any *other* ensemble. Furthermore, the cardinality of each ensemble is finite though very large, say $n!$, where n is the number of particles in each system and is of the order of 6×10^{23} . Therefore the union U of our ensembles will contain only finitely many systems. And hence it will be meaningful to make an assertion about a *majority* of the members of U .

We have been characterizing our ensembles with respect to two particular partitions, 3 and 4. But clearly, we can introduce at least a denumerable infinity of other partitions and can characterize our ensembles relatively to them, *mutatis mutandis*, in the same terms as we did with respect to 3 and 4. Hereafter we shall therefore not, in general, be restricted to any particular finite number of partitions.

The precise microstates E and F of our primary single system x were separated by some particular time interval Δt . We saw that during x 's transition from E to F in the time Δt , the direction of its entropy changes was partition-dependent: the entropy of x increased relatively to partition 3 while decreasing with respect to partition 4. Let us now consider each of the at least temporarily closed member-systems of our various ensembles $K_1, K_2, \dots, K_n$ at two different times t and $t + \Delta t$ in order to point out a fact of basic importance to our concerns: There is complete compatibility between the partition-dependence of the direction of the entropy changes undergone by x , on the one hand, and, on the other hand, the following partition-invariance of the temporal statistics of entropy change in the union U of our ensembles of systems: For any and every partition, the common initial entropy of all the systems in U will either have increased in a vast majority of the systems by the time $t + \Delta t$, or it will have remained the same.⁶ Note that we here used the indefinite article 'a' rather than the definite article 'the' when speaking of 'a vast majority of systems'.

For it is crucial to emphasize that the partition-invariance of an entropy increase or of entropy constancy is being asserted only in the sense that for every partition, *some majority or other* of the systems in U will exhibit the specified entropic behavior. Thus, the partition-invariance of the stated entropic behavior of a majority M of the systems in U does *not* at all require that the membership of the set constituting M relatively to a particular partition be itself partition-invariant! Evidently, the stated partition-invariance of the statistics of entropy change fully allows our finding above that during its transition from E to F , the single system x underwent specified oppositely directed entropy changes relatively to partitions 3 and 4. The latter finding does illustrate that the membership of M is NOT partition-invariant as follows: Relatively to partition 3, the particular system x does belong to M , whereas relatively to partition 4, x belongs to the minority of systems which decrease their entropy during Δt .

The claim that x thus belongs to what is only a minority of systems undergoing an entropy decrease relatively to partition 4 appears to be quite justified. The reason is that even within the single ensemble K_1 , which is included in U , in all likelihood *most* systems that started out in macrostate $D4E$ with entropy S_{D4E} , will *not* have evolved, after an interval Δt , into any distribution which is such that each cell of partition 4 con-

tains either exactly two particles or zero particles; nor into any other distribution possessing the lower entropy S_{D4F} of the latter kind of distribution; nor yet into any distribution whose entropy is lower than S_{D4E} . Relatively to partition 4, only a certain minority of systems decrease their initial entropy S_{D4E} during Δt , and α does belong to that minority with respect to this particular partition. But relatively to partition 3, a *different* minority of systems – though still only a minority! – decrease their initial entropy S_{D3E} during Δt , and α does *not* belong to that minority relatively to the latter partition; instead, with respect to partition 3, α increases its entropy. Thus, our partition-dependent findings concerning the system α accord entirely with the partition-invariance of the entropy statistics for the class U of systems. *Mutatis mutandis*, the same partition-invariance holds for the finitist entropy statistics in the respective unions of ensembles of *other* kinds of *like* closed systems.

Hence we draw the following very important conclusion: Although the direction of the entropy change exhibited by a single system such as α is indeed relative to the chosen partition, *no such relativity to the chosen partition characterizes the statistics of entropy change in the specified ensembles of systems*. For there is an invariance in the entropic behavior of a majority of the systems in U amid the relativity to the partition of the membership of that majority. I have called attention to the partition-invariance of the latter statistics of entropy change in order to discuss the bearing of the *physical significance* of this invariance on our central issue of anthropomorphism. Before we can turn to giving a statement of that bearing, we must deal more explicitly than above with the following triple status of the entropy: As we recall, it is the logarithm of a number W which is a measure of the degree of homogeneity as well as of the probability of occurrence, and not just the number of underlying microstates.

3. WHAT IS THE PHYSICAL SIGNIFICANCE OF THE TRIPLE ROLE OF THE ENTROPY FOR THE ENTROPY STATISTICS IN THE CLASS U ?

For the time being, our attention will be devoted to an individual physical system, and only later will we turn to our class U of systems.

(i) *Entropy as a Measure of the Degree of Homogeneity*

Perhaps it is possible to develop a notion of the degree of homogeneity or

equalization of a distribution which is *not* relativized to a partition of the phase space. At least in regard to the speeds of the particles, the possibility of such a measure is suggested by the fact that the Maxwell distribution of molecular velocities for the *equilibrium* state of uniform temperature gives a *partition-independent maximum spread* among the speeds of the particles, the spread being maximum within the confines of the fixed total energy of the system. And perhaps some function depending on the spatial distances or spatial spread among the n particles as well as on their velocity spread could be introduced to provide a partition-independent measure of homogeneity for the various macrostates of the system. To my knowledge, no such partition-independent measure of homogeneity has been worked out for 6-dimensional position-velocity space. As noted on p. 650, homogeneity is often called 'disorder'.

In any case, let us employ a concept of degree of homogeneity which is relativized to a partition. Hence we inquire whether the intuitively suggested, avowedly partition-dependent attribute of degree of homogeneity of a distribution D is rendered numerically by the quantity $W = n!/\pi n_i!$, *initially* construed as merely the number of different microstates belonging to D . After mentioning some obvious reasons for an affirmative answer to this question, I shall call attention to cases in which our intuition of *qualitative* relative homogeneity gives no clear-cut verdicts, in order to point out that I see no conflict between intuition and the homogeneity rankings furnished by the function W for these cases.

As for the obvious reasons, we observe that W is indeed maximum when the n particles are evenly allocated to the various cells, as compared to more or less *uneven* allocations among the cells of the given partition. Furthermore, two distributions, which differ only in that the sequence of particular, generally unequal numbers $n_1, n_2, \dots, n_m$ which specifies one of them is obtained by some permutation of the particular numbers specifying the other, clearly have the same degree of homogeneity on intuitive grounds. For intuitively, the degree of homogeneity does *not* depend on the *identity* of the particular cells in which given numbers of particles are distributed. And hence by assigning the same number to two distributions which differ in only this way, the function W implements our intuitive idea of relative degree of homogeneity in such cases.

But we can see that the function W also assigns the same number to certain distributions which differ far more strongly from one another. To

see this, note first that for a sufficiently large number of cells ($m \geq 4$), any given value of the product $\pi n_i!$ ($i = 1, 2, \dots, m$) resulting from one particular distribution is capable of various decompositions into factorials, such that $\sum_{i=1}^m n_i = n$ and such that the sequences of factorials in the different decompositions are *not* obtainable from one another by permutation. An arithmetically simple example of this kind of non-unique decomposition into *different* sets of m factorials is furnished by the physically uninteresting case of $n = 7$. For note that $24 = 4!1!1!1! = 3!2!2!0!$, while the condition $\sum_{i=1}^m n_i = n$ is satisfied, since $4 + 1 + 1 + 1 = 3 + 2 + 2 + 0 = 7$. Thus, the function W assigns the same degree of homogeneity even to some distributions which differ such that they are *not* obtainable from one another by moving the various *fixed* numbers of particles into other cells of the partition. Yet it seems to me that our qualitative intuition of degree of homogeneity does not interdict the verdicts of equihomogeneity furnished by W for distributions that differ from one another as just specified.

Thus, we can say that the numerical measures furnished by the function W render our *qualitative* intuitive idea that the relative degree of homogeneity of a distribution with respect to a given partition P is determined by *how evenly* the particles are distributed among the cells of P , the degree of evenness being a matter of how many different cells contain more or less nearly equal numbers of particles. And since we can speak of W and S interchangeably for our purposes, we can say that the entropy S provides a measure of the degree of homogeneity.

(ii) *Entropy as a Measure of the Probability of Occurrence of a Distribution*

Our concept of degree of homogeneity is relativized to a partition, and that degree is indeed measured by the entropy. Hence to say that a macrostate is a state of high entropy does tell us that it is a highly homogeneous distribution relatively to the partition with respect to which the given macrostate is both defined and of high entropy. But the assertion that high entropy states are states of great homogeneity does *not* thereby become a tautology. For to say that a macrostate is a distribution D of high entropy relatively to a partition P is to say much *more*, physically speaking, than that D is highly homogeneous relatively to P : It is to say *as well* that D has a high probability of occurrence in the time ensemble of macrostates which the given system attains relatively to the partition P . Hence the entropy is a property of D which has at least the following

objective macroscopic physical significance: It links or correlates the probability of the occurrence of a macrostate in time with its degree of homogeneity. I do not see how the physical significance of this important connection can be gainsaid, just because the entropy can *also* be regarded as a measure of the extent of human ignorance as to the actual microstate in the following sense: In any given occurrence of a certain macrostate D , only *one* of the W microstates that belong to D is actually being physically realized, but we don't know which one it is.

Hence I deny that W , as ascribed to the macrostate of the system at a given time, has a significance which is confined to the fact that, for all we know, *any one* of W different microstates is actually being realized at the time.

Even in its role as the mere number of different microstates belonging to a given kind of distribution D , the quantity W plays a physically relevant role and seems to me to be misleadingly belittled by being dubbed a measure of human ignorance. My reason for this claim will become clear after first noting the following: As is vouchsafed by the Birkhoff-von Neumann quasi-ergodic hypothesis, each one of the m^n microstates of which the permanently closed system is capable actually occurs *with the same frequency in time* or has the same probability $1/m^n$. And this ascription of equiprobability to all microstates is therefore *not* an assertion expressive of our lack of knowledge as to the actual (relative) frequency of the different microstates. In other words, having been justified by the quasi-ergodic hypothesis, this claim of equiprobability of microstates does *not* rest on an argument from *insufficient* reason, akin to the Laplacean's invocation of the principle of equi-ignorance (indifference) in defense of his *a priori* probability metric for dice. Let me therefore take *this* physical and epistemological status of the equiprobability of all microstates for granted. Then I can justify my contention above that even *qua* just being the number of different microstates belonging to a given distribution D , the quantity W is physically relevant, although only one of the W microstates belonging to D is being realized in any given occurrence of D . For note that since all microstates of which the closed system is capable actually occur equally often, the number W of microstates belonging to a given kind of distribution D *makes for* and determines the actual frequency of occurrence of that macrostate D in the time ensemble of the system's macrostates.

So much for the physical and epistemological credentials of the entropy as an attribute of the macrostates of an individual physical system. Let us now turn to our issue of anthropomorphism by using the results of our analysis to assess the *physical significance* of the statistics of entropy change in our class U . As will be recalled, the focus of our interest is whether the role of partition-dependence and thereby of human choice of the partition is such as to vitiate the physical significance of these statistics of entropy change and to relegate these statistics to being expressive of an anthropomorphism.

4. DO THE ROLES OF HUMAN DECISION AND IGNORANCE IMPUGN THE PHYSICAL SIGNIFICANCE OF THE ENTROPY STATISTICS FOR THE CLASS U ?

We noted the fundamental presumed fact that in the class U of systems, there is the following kind of partition-*invariance* of entropy increase or entropy constancy: For any and every partition, the common initial entropy of all the systems in U will either have increased in a vast majority M of the systems by the time $t + \Delta t$, or it will have remained the same. The substitution of one partition for another affects *only the membership of M* while leaving the specified partition-invariance of the entropy statistics intact. On this basis, we can now proceed to characterize several states of affairs as objective macroscopic physical facts, i.e., as *neither generated by a human choice of a partition nor expressive of human ignorance*. It is to be understood, of course, that the impending ascriptions of factual physical status are relative to the metrics of space and time ingredient in any of the partitionings of the 6-dimensional phase space and in the total volume and energy of each system.

It is a presumed fact that for any and every partition P , most systems in U which start out in a macrostate that is *comparatively inhomogeneous* relatively to P , after the time interval Δt will be in macrostates which are more homogeneous with respect to P as measured by W or S . The relativity of homogeneity itself to a partition P does not make it any less an objective physical fact that, by the time $t + \Delta t$, most members of the specified subclass of U will be in macrostates which *are* more homogeneous relatively to P . Moreover, these more or less homogeneous macrostates have the property of occurring more or less frequently in the respective

time ensembles of macrostates of their respective systems. And the fact that the lower and higher frequencies (probabilities) of occurrence pertain to distributions which are specified relatively to a partition does not detract from the objectivity of these frequencies. Nor are these various physical facts demoted to the status of anthropomorphisms just because they pertain to macrostates, and, as such, do not comprise the *still richer* factual physical content envisioned by *punctal* mechanics or fine-graining.

Thus, the statistics of entropy change in U codify partition-invariant macroscopic physical facts and cannot be held to be merely expressive of human ignorance of the underlying microprocesses. Human scientists do single out one or another particular partition relatively to which they characterize the macroscopic behavior of one or more systems entropically. But the partition-dependence of the direction of an entropy change in any one system no more makes the change of macrostate in that system anthropomorphic than the dependence on the inertial frame of the spatial distance between two events in Minkowski space-time renders that spatial distance anthropomorphic.

The gravamen of the charge of anthropomorphism which I have considered was that such partition-dependence as does obtain in the entropic descriptions of classical statistical mechanics renders these descriptions anthropomorphic. I have tried to argue that this charge is unfounded and that, at least to this extent, the use of the entropy concept is not illustrative of what J. L. Synge has aptly called the 'Pygmalion syndrome'.⁷

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the benefit of discussions with Allen Janis, Robert B. Griffiths, Joseph Camp and Richard Creath.

NOTES

* Originally written for *Festschrift for Henry Margenau* (ed. by E. Laszlo and E. B. Sellon), forthcoming.

¹ E. T. Jaynes, 'Gibbs vs. Boltzmann Entropies', *American Journal of Physics* **33** (1965) 398; italics in the original. Jaynes mentions in his Acknowledgments that he first heard the remark "Entropy is an anthropomorphic concept" from E. P. Wigner.

² Cf. E. T. Bell, *The Development of Mathematics*, McGraw-Hill Book Co., New York, 1945, 2nd ed., p. 170.

³ Cf. A. d'Abro, *The Rise of the New Physics*, Dover, New York, 1951, pp. 908-909.

⁴ The fact that $W_0 \cdot d_0$ microstates rather than only W_0 of them underlie an entropy state

S_0 was erroneously overlooked on p. 162 of my paper 'The Anisotropy of Time', in T. Gold (ed.), *The Nature of Time*, Cornell University Press, 1966. Hence the statement of boundary conditions governing the so-called branch systems given there must be amended: Either the random samples of the formulation must be particularized to each distribution of given entropy, as I shall do later on in this Section 2, *or* the random samples must be taken from the *total* number ($W_1 \cdot d_1$) of microstates which can underlie an entropy state S_1 of the given kind of system.

A like correction applies to pp. 256–257 in Chapter 8 of the present volume.

⁵ James and James, *Mathematics Dictionary*, 3rd edition, D. Van Nostrand Co., N.J., 1968, p. 300.

⁶ For a statement of the physical reasons underlying this statistical claim as such, see pp. 256–258 of Chapter 8.

⁷ J. L. Synge, *Talking About Relativity*, North-Holland Publishing Co., Amsterdam and London, 1970, p. 8.

A PANEL DISCUSSION OF SIMULTANEITY BY SLOW CLOCK TRANSPORT IN THE SPECIAL AND GENERAL THEORIES OF RELATIVITY

INTRODUCTION: THE CONTEXT OF THESE ESSAYS*

In his famous 1905 paper on the special theory of relativity (STR), before the subject of the *relativity* of simultaneity in systems in relative motion is even broached, Einstein enunciates explicitly his doctrine of the definitional character of simultaneity in a single “stationary” system ([2], p. 40). Subsequent interpreters, most notably Reichenbach, have maintained that Einstein was correctly claiming the relation of simultaneity in a single inertial system to be conventional in a significant and nontrivial sense ([4], p. 127). Einstein there describes a method of synchronizing clocks located at different places in the stationary system by a signalling process. Suppose two clocks U_A and U_B are located at places A and B respectively. Let a light signal be sent from A to B, where it is immediately reflected back to A. Let t_1 be the time at which the signal departs from A, and let t_3 be the time of its return to A after reflection at B, both times measured on U_A . Einstein establishes a synchrony between U_A and U_B by *stipulating* that the time taken for the signal to travel from A to B equals the time it takes to return from B to A. As Reichenbach expressed it, a relation of simultaneity is established by the conventional choice of ϵ in the formula

$$t_2 = t_1 + \epsilon(t_3 - t_1),$$

where t_2 is the time of arrival of the light signal at B, and subject only to the condition

$$0 < \epsilon < 1.$$

Append
§§ 57–59
and 60

* This Introduction on pp. 666–669 was co-authored with Wesley C. Salmon and is included in this chapter by his permission.

Einstein's stipulation that the to and fro signal velocities are equal to one another in magnitude is tantamount to the choice $\epsilon = 1/2$. According to Reichenbach, Einstein's choice is distinguished by its vast *descriptive* simplicity, but no other choice would involve any factual error. The restriction of ϵ to the open interval $(0, 1)$ is designed to secure the topological result that the signal's arrival at B is always between its departure and return on any given trip.

Two important physical considerations provide the foundation for the claim that distant simultaneity is nontrivially conventional:

- (1) Clocks originally synchronized with one another at some place and then transported along different paths or at different velocities are, in general, found to be out of synchrony when brought together again.
- (2) There is a finite maximum round-trip velocity of transmission of signals.

On the basis of these two fundamental facts, it has been concluded that neither clock transport nor the sending of signals is sufficient to determine a relation of metrical simultaneity. Hence, it has been argued, a convention is required, and Einstein fulfilled this need with his stipulation of equal travel times for light in the opposite directions of the round trip.

The reason that *actual* transported clocks do not spontaneously determine a relation of simultaneity is clear from (1) above. If, for instance, we have three clocks, C_1, C_2, C_3 , all synchronized at place A, and if we move C_2 and C_3 to a place B along the same path at different velocities, C_2 and C_3 will no longer be synchronized when both are located at B. Thus, if we take the *uncorrected* reading of a transported clock as determining the relation of simultaneity between events at A and B, we do not get a unique relation but inconsistent ones. For the relation depends upon the mode of transport of the particular clock that is chosen. Moreover, if either clock, say C_3 , is returned to A, it will be found to be out of synchrony with C_1 . These considerations evidently led Einstein to abandon the transported clock as a basis for the simultaneity relation, and to adopt instead the concept of simultaneity based upon light signals. Once a criterion of distant simultaneity is given, it is possible to describe quantitatively the effect of a trip at a certain velocity upon a transported clock. For Einstein's signal synchrony the results are expressed in the well-known time dilation equations.

In his posthumous book on the theory of relativity, P. W. Bridgman maintained that one can dispense with Einstein's light signals as a means of synchronizing clocks in the inertial systems of the special theory of relativity ([1], pp. 64–67). Bridgman first *adjacently* synchronizes each of a number of clocks $C_1, C_2, \dots, C_n$ with a stationary clock U_A at essentially the same space point A. And then he transports the clocks C_n to another space point B along AB such that they do not arrive jointly. Thereupon he sets the clock U_B at B once as follows: he eliminates any difference k existing in the limit of infinitely slow transport of clocks C_n between their readings and those of the stationary clock U_B .

One fundamental difficulty might seem to beset Bridgman's procedure. In order to describe the method of slow transport synchronization it is necessary to refer to the velocities at which the clocks are transported from A to B, but this velocity

itself seems to make sense only in the presence of a simultaneity relation that enables us to speak of related times at different places. How else can we say how much time it took for clock C_i to get from A to B? Bridgman deals with this problem by using the "self-measured velocity" as read off on the clock C_i itself. As these "velocities" go to zero, he observes, we get a uniquely determined simultaneity relation ([1], p. 65).

Bridgman points out that if the STR is true, then his method yields the clock readings of the Lorentz transformations no less uniquely than does Einstein's standard procedure of synchronization by light signals. Also, the method's reliance on extrapolation to the limit of infinitely slow clock transport does not prevent it from being carried out in an acceptably short time. In the sequel, the term "slow clock transport" is to be understood as an abbreviation for "infinitely slow clock transport." Bridgman goes on to comment on the philosophical significance of the feasibility of synchronism by slow clock transport as an alternative to the reliance on light signals. Specifically, he assesses Einstein's conception of simultaneity as articulated by Reichenbach, writing:

Distant clocks set by the transport method will agree with clocks set by Einstein's method. . . . the transported clock . . . will give a value for Reichenbach's ϵ equal to $\frac{1}{2}$ What is the significance of this? Does it mean that there is something "absolute" about the value $\frac{1}{2}$? What becomes of Einstein's insistence that his method for setting distant clocks—that is, choosing the value $\frac{1}{2}$ for ϵ —constituted a "definition" of distant simultaneity? It seems to me that Einstein's remark is by no means invalidated. He was saying, in effect, that any method whatever for setting distant clocks involves an element of definition, and that in choosing the value $\frac{1}{2}$ he was merely adopting a particular one of these methods ([1], p. 66).

. . . the decision to use one or the other method is a decision in our control, involving a corresponding *definition* of distant simultaneity. The fact that these two methods agree . . . is in no wise a logically necessary fact, but is something that has to be established by independent experiment ([1], p. 67).

In 1967, B. Ellis and P. Bowman published a modified version of Bridgman's method of synchronism by slow clock transport ([3]; hereafter, this paper will be cited as "*E & B*"). These authors call attention to those physical facts of clock behavior which underlie the method. And they contend that these physical facts refute Einstein's philosophical conception of metrical simultaneity in the STR as interpreted by Reichenbach. Ellis and Bowman make a point of the physical merit of Bridgman's idea of synchronization by slow transport. But they do not indicate any awareness of differing from him as to the import of his method for the *philosophical* validity of Einstein's conception of simultaneity as articulated by Reichenbach. Unlike Bridgman, Ellis and Bowman maintain that the physical presuppositions of his method demonstrate either the falsity or the triviality of the thesis that in the STR, distant simultaneity involves an important conventional ingredient.

Consider two spatially separated clocks at rest in an inertial system K. Ellis and Bowman speak of such clocks as being "in standard synchrony" in K, if their synchronism is the one resulting alike from Einstein's light signal rule or from slow clock transport, and they speak of nonstandard synchrony in the case of clock settings which correspond to a value of $\epsilon \neq \frac{1}{2}$. They summarize the claims of their paper as follows (*E & B*, p. 116):

It has often been claimed that there are no logical or physical reasons for preferring standard signal synchronizations to any of a range of possible nonstandard ones. In this paper, the range of consistent nonstandard signal synchronizations, first for any one inertial system, and second for any set of such systems, is investigated, and it is shown that the requirement of consistency leaves much less room for choice than is commonly supposed. Nevertheless consistent nonstandard signal synchronizations appear to be possible. However, it is also shown that good physical reasons for preferring standard signal synchronizations exist, if the Special Theory of Relativity yields correct predictions.

The thesis of the conventionality of distant simultaneity espoused particularly by Reichenbach and Grünbaum is thus either trivialized or refuted.

In general conformity to the order of *E* & *B*'s paper, the issues raised by them directly or indirectly will be discussed in the succeeding four papers by A. Grünbaum, W. C. Salmon, B. van Fraassen and A. Janis. But Janis's paper extends the discussion to the feasibility of synchronism by slow transport in the *non*-inertial frames of the General Theory of Relativity (GTR). None of the four authors disputes the claim that in the inertial frames of the STR, the method of slow transport yields a unique synchrony which coincides with $\epsilon = 1/2$.

The four papers which are to follow grew out of an active exchange of views and meeting of minds among these four authors. Despite this give-and-take, it was felt useful to preserve individual authorship to the extent of not merging their essays. But each of the four authors has had the benefit of the comments of at least one of the others. And philosophically, their essays constitute a single whole.

REFERENCES

- [1] Bridgman, P. W., *A Sophisticate's Primer of Relativity*, Middletown, Wesleyan University Press, 1962.
- [2] Einstein, A., "On the Electrodynamics of Moving Bodies," *The Principle of Relativity: A Collection of Original Memoirs*, New York, Dover Publications, Inc., 1952.
- [3] Ellis, B. and Bowman, P., "Conventionality in Distant Simultaneity," *Philosophy of Science*, vol. 34, 1967, pp. 116-136.
- [4] Reichenbach, H., *The Philosophy of Space and Time*, New York, Dover Publications, 1958.

1. Summary. Ellis and Bowman's account of nonstandard signal synchronizations is examined as a prolegomenon to this paper. Attention is called to some consequences of an important ambiguity in their account of the *transitivity* of nonstandard synchrony. Then an analysis is given of the principle of relativity (first postulate of the STR) to assess *E* & *B*'s claim that this principle either restricts nonstandard signal synchronisms or rules them out altogether. It is argued that the latitude for choices of nonstandard synchronisms is *not* circumscribed by the factual content of the principle of relativity; instead, the exclusion of such synchronisms by this principle depends on a tacit appeal to the *particular* conventions implicit in certain formulations of the principle. Hence *E* & *B*'s claim is rejected as an argument against the factual tenability of nonstandard synchronisms.

The second part of this paper offers a detailed critique of *E* & *B*'s central philosophical thesis. It is contended that in the STR, synchronism by slow clock transport neither refutes nor trivializes the ingreience of a convention in that theory's distant simultaneity. To provide a basis for this defense of the Einstein-Reichenbach conception of simultaneity, attention is first given to the bearing of clock transport and of causal chains generally, including gravitational influences, on simultaneity in Newton's world. It is pointed out how these Newtonian physical agencies make for simultaneity relations which are *both* intersystemically invariant *and* nonconventional. Thus, the absolute simultaneity relations in Newton's world are shown to hold as a matter of *fact*, although the particular identity of the time coordinate assigned alike to all members of a simultaneity class is trivially conventional. Here it is also noted why there is no incompatibility in Newton's world between the simultaneity of two events and their being the termini of gravitational influence chains, whereas *all* causally connectible events are nonsimultaneous in the universe of the STR.

Next, the status of simultaneity in a so-called "quasi-Newtonian" world is assessed. The latter world differs from Newton's as follows: both light rays and gravitational influences have the same *finite* round-trip velocity in any one inertial system, while also being the fastest causal chains capable of connecting any two space points. This world resembles the Newtonian one in that transported clocks can be said to furnish *consistent* and even invariant simultaneity relations. But, in opposition to *E* & *B*, it is argued that in the quasi-Newtonian world, simultaneity relations involve a *non-trivial* conventional ingredient, in contrast to those of Newton's world, which hold as a matter of temporal physical fact.

Finally, it is indeed agreed that the slowly transported clock *presents us* with a

* The author wishes to thank the National Science Foundation for the support of research.

unique time coordinate for any particular event, while the light ray of signal synchrony delivers no such time coordinate except by our stipulation. But it is held that this physical fact does not sustain *E* & *B*'s philosophical thesis. Comparison of simultaneity by slow clock transport synchrony in the universe of the STR with simultaneity in the quasi-Newtonian world refutes the philosophical thesis which *E* & *B* have based on the facts of slow clock transport. For this comparison shows that in the STR, slow clock transport cannot confer *factual* physical truth on the particular simultaneity relations asserted by the Lorentz transformations, any more than an exchange of light signals can do so. The Einstein–Reichenbach conception of simultaneity is thus vindicated in the face of the facts of slow clock transport.

This conclusion will be reinforced by the results from the GTR in A. Janis's paper below. Janis shows that there is a large class of non-inertial frames, which includes the rotating disk in flat space-time, in which slow transport synchrony fails. Hence there would be scope for conventional choices of simultaneity in all of these cases, even if *E* & *B* had given a correct philosophical characterization of simultaneity by slow transport synchrony.

2. Examination of *E* & *B*'s Account of Nonstandard Signal Synchronizations. Let t_1 be the time on the clock U_A at the point *A* at which a light ray is emitted directly toward the point *B*, and t_3 the time on U_A at which this light ray $\overline{ABA}$ returns directly to *A* after being instantaneously reflected at *B*. Let the distance between *A* and *B* be d . We are given the empirical principle of the STR that the *round-trip* or *two-way* velocity of light $2d/(t_3 - t_1)$ which is measured at *A* has the numerical value c *in vacuo*. *E* & *B* refer to this principle as “the two-way light principle.” Let us now impose the condition—which *E* & *B* (p. 117) call “the topological condition for light”—that the time number t_2 which the clock U_B is set to read when the light-ray reaches *B* is to be *between* t_1 and t_3 . Using Reichenbach's notation, we have

$$(1) \quad t_2 = t_1 + \epsilon(t_3 - t_1), \quad \text{where } 0 < \epsilon < 1.$$

To emphasize that in this equation the particular value of ϵ specifies the particular setting of the clock at *B* required to qualify it as synchronized with the one at *A*, we shall write ϵ_{AB} . And we shall assert that U_B is synchronized with U_A according to that particular value ϵ_{AB} by writing:

$$U_B \text{ syn}(\epsilon_{AB}) U_A$$

Let it be granted as an empirical fact, implicit in the two-way light principle, that the direct light ray which departs from *B* at the above U_B time t_2 is instantaneously reflected at *A* and returns directly to *B* at a time t_4 on U_B which satisfies the following principle of isotropy of *round-trip* times:

$$(2) \quad t_4 - t_2 = t_3 - t_1.$$

Let us denote the latter round trip time by T . It is clear that the principles and conditions laid down thus far do *not* entail $\epsilon_{AB} = \epsilon_{BA}$.

But it should likewise be clear that none of our commitments thus far *require* us

to accept the reading t_3 of U_A on the arrival of the light ray as *automatically* qualifying U_A to be synchronized from the point of view of U_B or of some third clock U_C . It is true that U_A 's reading t_3 (along with its reading t_1) was already invoked to impart a setting t_2 to U_B which satisfies the condition

$$U_B \text{ syn}(\epsilon_{AB})U_A.$$

And the principles asserted thus far permit us to *tamper* with the otherwise existing setting of U_B in order to satisfy the latter condition of synchronism. But these principles also allow us to "*correct*" computationally, if necessary, the reading t_3 of U_A to assure the fulfillment of the condition

$$U_A \text{ syn}(\epsilon_{BA})U_B$$

for any chosen ϵ_{BA} between 0 and 1, or the fulfillment of the corresponding condition

$$U_A \text{ syn}(\epsilon_{CA})U_C$$

for a clock U_C whose setting is unspecified. It is to be understood that computational correcting of the reading t_3 of U_A does *not* involve physically tampering with the setting of U_A . If the latter kind of "correcting" is to be disallowed, we must introduce the following restriction as governing any clock X whose readings have already been used to synchronize some other clock Y with X via some ϵ_{XY} between 0 and 1: the setting of any such clock X *must* be accepted as automatically qualifying X to be synchronized from the point of view of Y and of any other clock Z . This restriction will be called *the rule of committed synchronism*, and abbreviated to "**RCS**."

If we do accept **RCS** and hence the t_3 reading of U_A , then, of course, the value ϵ_{BA} is uniquely determined by the equation

$$(3) \quad t_3 = t_2 + \epsilon_{BA}T.$$

And having used **RCS** to obtain (3), we can use (1) and (3) to make substitutions for $t_2 - t_1$ and $t_3 - t_2$ respectively in the arithmetical identity

$$(4) \quad t_2 - t_1 + t_3 - t_2 = t_3 - t_1 = T,$$

obtaining $\epsilon_{AB}T + \epsilon_{BA}T = T$, or

$$(5) \quad \epsilon_{AB} + \epsilon_{BA} = 1.$$

E & *B* assert erroneously (p. 117) that (5) follows from the two-way light principle. They define ϵ_{AB} as "the one-way velocity of light" from A to B, i.e. they define $\epsilon_{AB} \equiv d_i(t_2 - t_1)$. Here the reading t_2 is held to qualify U_B as synchronous with U_A . The two-way light principle tells us that $T = 2d/c$. But the conjunction of the arithmetical identity (4) with this principle cannot yield

$$(5a) \quad \frac{d}{\epsilon_{AB}} + \frac{d}{\epsilon_{BA}} = \frac{2d}{c},$$

unless **RCS** is invoked to *accept* the reading t_3 as qualifying U_A to be synchronous with U_B and write

$$(5b) \quad \epsilon_{BA} = \frac{d}{t_3 - t_2}.$$

Yet E & B use (5a) to deduce (5) via the relations

$$\epsilon_{AB} = \frac{c}{2c_{AB}} \quad \text{and} \quad \epsilon_{BA} = \frac{c}{2c_{BA}}.$$

Since the *two-way* light principle cannot itself entail any relation between the two distinct *one-way* synchronisms specified by ϵ_{AB} and ϵ_{BA} , the two-way principle cannot entail (5).

E & B 's tacit use of **RCS** would be a matter of small importance if not for the following consequential fact: E & B 's account of nonstandard signal synchrony suffers from an ambiguity between a *generalized* form of **RCS** (hereafter called "**GRCS**") and a much stronger principle which they call "the transitivity of synchrony" (p. 117). **GRCS** is the rule that at most one act of setting is permissible per clock to achieve the *mutual* synchronism of any two or more clocks. Thus, if a clock Y has been synchronized from the point of view of a clock X , then the resulting setting of Y must be accepted as a basis for synchronizing any other clock Z with Y , and the setting of a clock Z resulting from using Y to synchronize Z must automatically be accepted as rendering Z synchronous with X . E & B first mention the *transitivity* of synchrony by reference to the *standard* criterion of synchrony ($\epsilon = 1/2$) before they "enquire how else clocks may be set, and yet be said to be synchronous" (p. 117). And immediately prior to formulating the *standard* criterion, they state the transitivity of synchrony in the context of that criterion as follows (p. 117):

See Append.
§57

See Append.
§54

... for any three points A, B, and C at rest in any inertial system K, if clocks at A and B are synchronous according to the following criterion, and clocks at B and C are synchronous according to the same criterion, then the clocks at A and C are synchronous according to this criterion. We call this principle that of the *transitivity of synchrony* and the criterion that for *standard signal synchrony*.

Note that in the principle of transitivity the second of the three pairs of clocks is explicitly required to be synchronized "according to the *same* criterion" (italics not in the original) as the first, and that the principle then guarantees the automatic synchronism of the third pair of clocks "according to this criterion." By contrast, **GRCS** in no way requires that the second pair of clocks be synchronized according to the same criterion as the first. **GRCS** does *decree* the synchronism of the *third* pair of clocks via the two synchronizations of the first two pairs. But, unlike the principle of transitivity, **GRCS** does not claim that this decreed synchronism of the third pair conforms "to the same criterion" as the synchronism of the first two pairs.

The writings of Reichenbach figure prominently in the literature on signal synchronization to which E & B address themselves. And E & B follow Reichenbach in formulating the principle of transitivity of synchronism with respect to a *single* criterion. Reichenbach speaks of the latter as being "the *same rule*" of synchronization ([14], p. 168; his italics). For him, a single criterion or rule of synchronism is constituted by the use of one particular numerical value of ϵ .¹ And hence transitivity has been understood as restricting the three values ϵ_{AB} , ϵ_{BC} , and ϵ_{AC} to being *equal*. At any rate, let us first note some of the consequences of *this construal*

¹ See, for example, Reichenbach's treatment of clock synchronism on a rotating disk in his [13], pp. 140–141.

of transitivity, which requires all three values of ϵ to be equal. Then we note that, contrariwise, **GRCS** neither restricts the first two of these values to being equal, nor asserts that if they were equal, then ϵ_{AC} would likewise be equal to them.

It will be useful to illustrate the difference between the mere fulfillment of **GRCS** and the obtaining of transitivity by reference to values of ϵ_{AB} , ϵ_{BC} , and ϵ_{AC} which satisfy the condition $0 < \epsilon < 1$. We shall do so by showing that **GRCS** can readily be fulfilled for suitable triplets of such values $\epsilon \neq \frac{1}{2}$, while *all non-standard* synchronisms $0 < \epsilon < 1$ are *intransitive*.²

See Appendix §55 *Proof of the Intransitivity of Nonstandard Synchronisms for $0 < \epsilon < 1$.*

Consider three clocks, A , B , and C stationed at corresponding points of an inertial system. Let two light rays be jointly emitted from A at $t = 0$ as follows: one light ray reaches C directly after traversing the straight line segment AC whose length is n , while the other ray first traverses the segment AB of length l and then the segment BC of length m , arriving at C later than the first ray.

The time-increment on the clock C between the arrivals of the two rays at C is *independent* of how the clocks A , B , and C are synchronized. Therefore, we can calculate that time increment on C from information which is formulated in terms of the standard synchronism $\epsilon = \frac{1}{2}$ without prejudice to our impending inquiry into the transitivity of a *non-standard* synchronism $0 < \epsilon < 1$.

In the case of the standard synchronism, the direct light ray $\overline{AC}$ reaches C when the C -clock reads $t_C^1 = n/c$, while the circuitous light ray $\overline{ABC}$ arrives at C at the later time $t_C^2 = l/c + m/c$, where c is the usual speed of light. Hence the time increment Δt_C on the C -clock between the arrivals of the two rays is

$$(6) \quad \Delta t_C = \frac{l}{c} + \frac{m}{c} - \frac{n}{c}.$$

If there is to be transitivity with respect to any *non-standard* synchronism $\epsilon_{AB} = \epsilon_{BC} = \epsilon_{AC}$ between 0 and 1, the facts of light propagation must vouchsafe for such ϵ that the following conditional statement is true: Given

$$C \text{ syn } (\epsilon_{BC})B, \quad \text{and} \quad B \text{ syn } (\epsilon_{AB})A,$$

then

$$C \text{ syn } (\epsilon_{AC})A.$$

Let T_l , T_m , and T_n denote the respective *round-trip* times on the clocks A , B , and C for light rays $\overline{ABA}$, $\overline{BCB}$, and $\overline{ACA}$. By the two-way light principle, we have

$$(7) \quad T_l = \frac{2l}{c}, \quad T_m = \frac{2m}{c}, \quad \text{and} \quad T_n = \frac{2n}{c}.$$

After departing from A at $t = 0$, the direct light ray reaches C at the time $\tau_1 = \epsilon_{AC}T_n$,

² Intransitivity of a relation R is a stronger property than non-transitivity. For a relation R is *intransitive* in the set S if, for any three elements x , y , and z of S , the formulas $x R y$ and $y R z$ imply the formula $\sim(x R z)$. But a relation is *non-transitive* in S , if the first two formulas merely do *not* imply $x R z$.

whereas the circuitous ray arrives at B at the time $t_B = \epsilon_{AB}T_i$, and then reaches C at the time

$$\tau_2 = \epsilon_{AB}T_i + \epsilon_{BC}T_m.$$

To implement the requirement **GRCS** with respect to clock C, it suffices that ϵ_{AB} , ϵ_{BC} , and ϵ_{AC} are so chosen that at most one setting of any one clock is needed to satisfy the condition

$$(8) \quad \tau_1 + \Delta t_C = \tau_2.$$

Using (6) and (7), this equation can be written in the form

$$(9) \quad \frac{2n}{c} \epsilon_{AC} + \frac{l + m - n}{c} = \frac{2}{c} (l\epsilon_{AB} + m\epsilon_{BC}),$$

which yields

$$(10) \quad \frac{l + m - n}{2} = l\epsilon_{AB} + m\epsilon_{BC} - n\epsilon_{AC}.$$

The specified fulfillment of condition (10) implements the requirement **GRCS** for clock C. But this same condition (10) will now enable us to see that even though there are nonstandard synchronisms $0 < \epsilon < 1$ which thus fulfill **GRCS**, *any such synchronism is intransitive*.

To prove the latter intransitivity, we first determine what is entailed by (10) upon imposing the transitivity requirement

$$\epsilon_{AB} = \epsilon_{BC} = \epsilon_{AC} = \epsilon.$$

In that case, (10) becomes

$$\frac{1}{2} (l + m - n) = \epsilon(l + m - n)$$

or, since $l + m - n \neq 0$,

$$\epsilon = \frac{1}{2}.$$

Thus we have shown that transitivity of synchronism *entails* the *standard* synchronism $\epsilon = \frac{1}{2}$. And therefore we have here furnished a proof that standard synchronism is a *necessary* condition for the transitivity of synchronism. This result supplements Reichenbach's proof ([13], pp. 34–38, esp. p. 38) of Einstein's statement that $\epsilon = \frac{1}{2}$ is *sufficient* for this kind of transitivity of synchronism ([3], p. 40).

This much establishes the *non-transitivity* of nonstandard synchronisms $0 < \epsilon < 1$. To prove their *intransitivity*, we use equation (10) to show that if $\epsilon_{AB} = \epsilon_{BC} = \epsilon_k$ such that $\epsilon_k \neq \frac{1}{2}$, then $\epsilon_{AC} \neq \epsilon_k$. Under the hypothesis of $\epsilon_{AB} = \epsilon_{BC} = \epsilon_k$, we can write (10) in the form

$$(10') \quad (l + m)(\frac{1}{2} - \epsilon_k) = n(\frac{1}{2} - \epsilon_{AC}).$$

Since l , m , and n are the respective lengths of the sides of the Euclidean triangle ABC, we are given that

$$(l + m) \neq n,$$

while *neither* side of this inequality is zero! Hence equation (10') can hold only under one of the following two conditions:

$$(i) \quad \frac{1}{2} - \epsilon_k = \frac{1}{2} - \epsilon_{AC} = 0,$$

which is disallowed by the hypothesis of *non*-standard synchronization $\epsilon_k \neq \frac{1}{2}$ or

$$(ii) \quad \frac{1}{2} - \epsilon_k \neq \frac{1}{2} - \epsilon_{AC}.$$

Hence $\epsilon_k \neq \epsilon_{AC}$, and the nonstandard synchronism $\epsilon_k \neq \frac{1}{2}$ is intransitive. **Q.E.D.**

To illustrate that there are nonstandard synchronisms $0 < \epsilon < 1$ which fulfill **GRCS** for clock *C*, we select the arithmetically simple case of applying (10) to an equilateral triangle *ABC*, so that $l = m = n$. And we merely note that the argument can be generalized to the arithmetically complicated case of a *scalene* triangle. In the equilateral case, (10) becomes

$$(11) \quad \epsilon_{AC} = \epsilon_{AB} + \epsilon_{BC} - \frac{1}{2}.$$

And (11) tells us, for example, that if $\epsilon_{AB} = \epsilon_{BC} = \frac{1}{10}$, then $\epsilon_{AC} = \frac{3}{10}$. Moreover, (11) shows us that **GRCS** can be satisfied for clock *C* by equal values of ϵ_{AB} and ϵ_{BC} between 0 and 1, coupled with values of ϵ_{AC} that do *not* belong to that interval. For example, (11) yields the values given in the following tabulation.

ϵ_{AB}	ϵ_{BC}	ϵ_{AC}
$\frac{1}{8}$	$\frac{1}{8}$	$-\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$	0
$\frac{3}{4}$	$\frac{3}{4}$	1
$\frac{9}{10}$	$\frac{9}{10}$	$\frac{13}{10}$

But these values of ϵ_{AC} do not satisfy *E* & *B*'s topological condition of light $0 < \epsilon < 1$. And when laying down the latter condition, *E* & *B* make clear that it must be met by any two clocks *A* and *B* of which they would "be prepared to say that the clocks at *A* and *B* were synchronous" (p. 117).

The results which we just deduced from equation (10) yield the following conclusions:

(i) In the single criterion sense of "transitive" synchronism defined by *E* & *B*'s "principle of the transitivity of synchrony" (p. 117), standard synchronism is the only kind of transitive synchrony.

(ii) **GRCS** is satisfied not only by the intransitive nonstandard synchronisms that fulfill the topological condition for light but also by some of those which violate it.

These conclusions now enable us to evaluate the following statement by *E* & *B* (pp. 118-119):

Now, from the two-way light principle, together with that of the transitivity of synchrony, we can easily derive a *three-way light principle*. . . it is obvious that, by successive applications of the three-way light principle, a general *n-way light principle* can be derived. . . We require that any nonstandard light signal synchronization should satisfy the *n-way light principle*. For if this principle were not satisfied, *the light signal relationships between*

the clocks of our system would not be transitive, and hence would not define a relationship of synchrony. [Italics in this sentence not in the original.]

The simplest kind of nonstandard signal synchronization is one in which ϵ_{AB} is chosen to be a function only of the direction from A to B, and not of the position of A or of the distance of B from A. It is, moreover, such that the variation of ϵ with direction is symmetrical about some chosen axis, say the X -axis of a rectangular coordinate system. In this kind of nonstandard signal synchronization, ϵ is the same for all directed lines inclined at an angle θ to the positive direction of the X -axis, and may be designated by ϵ_θ . The value of ϵ_θ can now be calculated using the three-way light principle.

Here E & B tell us that “any nonstandard light signal synchronization” must be transitive in order to qualify as “a relationship of synchrony.” But if this were so, then it would follow from our results above that there could be no nonstandard light signal synchronization at all! And in that case, the formula for nonstandard ϵ_θ which E & B go on to deduce (p. 120) would have no relevance at all. But their deduction of the n -way light principle and of the formula for ϵ_θ actually employs not, as they claim, the principle of transitivity but merely the weaker requirement that we called **GRCS**. For E & B invoke “transitivity” (p. 118) only to justify an equation which is equivalent to our **GRCS** equation (8) above.

Wesley Salmon has pointed out that whatever may have been the construal of being a “single criterion” of synchronism in the prior literature on the transitivity of that relation, it is unreasonable to construe a single rule of synchronism restrictively as requiring that the same *numerical* value of ϵ be used. In Salmon’s view, one should countenance E & B ’s formula for ϵ_θ as a single rule. On this broader construal of transitivity, my proof that *standard* synchronism is *uniquely* transitive does not, of course, apply, and the distinction between **GRCS** and the principle of transitivity disappears.

Having used **GRCS** to derive their formula for ϵ_θ , which they call “the distribution law for light velocities” (p. 121), E & B write (p. 121):

We conclude that at least one kind of nonstandard signal synchronization of the clocks of any inertial system is possible. . . . It is possible that other kinds of nonstandard signal synchronization are possible. But these have never been considered in the literature, and those who have spoken of the conventionality of clock synchronization procedures have always had in mind nonstandard synchronizations of the kind we have analyzed.

E & B then proceed to appeal to the distribution law for light velocities, i.e. *their* formula for ϵ_θ , to assess the acceptability of a non-standard kind of synchronism proposed by Reichenbach ([14], pp. 162–164). But the kind of non-standard synchronizations which E & B analyzed were characterized by the requirement **GRCS** (as well as by the topological condition for light). And Reichenbach had disavowed this requirement in his account of nonstandard synchronizations. For when discussing the case in which arbitrary ϵ between 0 and 1 are used to synchronize two or more clocks with one particular central master clock, Reichenbach said explicitly: “For the comparison of any given clocks a special value of ϵ , *which may vary with time*, will have to be chosen” (italics not in the original) ([14], p. 168). Since Reichenbach had rejected **GRCS**, E & B are mistaken in claiming (p. 121) that:

. . . those who have spoken of the conventionality of clock synchronization procedures have always had in mind nonstandard synchronizations of the kind we have analysed.

See Append.
§56

And since E & B 's distribution law of light velocities is predicated on GRCS, it is irrelevant for them to indict Reichenbach's proposed nonstandard signal synchronization as "demonstrably inconsistent" with that law (p. 121, n. 3). Besides, even if that indictment were relevant, E & B did not establish that "Reichenbach's proposed nonstandard signal synchronization is therefore intransitive and hence unacceptable." For the violation of GRCS by Reichenbach's synchronization shows only that the latter is *non-transitive* but not that it has the stronger property of intransitivity.

Non-Standard Synchronisms and the Principle of Relativity.

Before arguing for their philosophical interpretation of synchronism by slow clock transport, E & B (p. 122) turn next to restrictions which they believe to be imposed on nonstandard signal synchronizations by the principle of relativity (first postulate of the STR). They consider both a weak and a strong version of this principle. And they link these versions to Einstein and Poincaré respectively. E & B are concerned to point out that (i) the weak version disallows a very large proper subclass of the nonstandard signal synchronizations $0 < \epsilon < 1$, and (ii) if we require that Newton's *first* law of motion holds in every inertial system, then the strong version of the principle of relativity rules out *all* nonstandard signal synchronizations. They then argue that these presumed violations of the principle of relativity by nonstandard synchronisms $\epsilon \neq 1/2$ show the following: The analogy in a statement of mine which they cite (p. 122, n. 5) has been pushed too far. I had stated that "the use of *different* values of ϵ in different inertial systems S and S' can no more violate the First Postulate of the STR than the use of centimeter units to express the velocity of light in S and of miles to express that same velocity in S' , . . ." ([5], p. 191).

E & B characterize the "weak or Einstein version" and "the strong or Poincaré version" of the principle of relativity respectively as follows (p. 122, n. 6):

In its usual statement it asserts the sameness, equivalence, or invariance of laws with respect to inertial frames in uniform rectilinear motion . . . and this sameness is often taken as referring to the expression of those laws *under the Lorentz transformations*. This version is usually associated with Einstein. . . .

Another version of the relativity principle asserts the complete impossibility of determining absolute motion between inertial frames. . . . This version is stronger than the first mentioned insofar as it also rules out differences in quantitative determinations (say, of length, velocity, permittivity, etc.) between different inertial systems.

Now "assuming Grünbaum espouses the weak or Einstein version" (p. 122, n. 5), E & B believe my analogy to units to be refuted by the fact that "the choice of length scale (within the same dimension) does not affect the form of the expression of any laws of physics, and thus does not violate the principle of relativity, while the choice of $\epsilon_0 \neq 1/2$ obviously would." And they conclude (p. 125, n. 8) that the quoted "Grünbaum claim . . . is mistaken even if he espouses the strong relativity principle." They rest the *latter* conclusion on the following three premisses: (i) a result of McPhee's (pp. 124–125):

If we synchronize the clocks of various inertial systems so as to ensure that a uniform straight line motion in any one system is seen as a uniform straight line motion in any other,

and that the condition of the reciprocity of relative velocities is satisfied, then the clocks of each inertial system will be found to be in standard signal synchrony.

By the condition of reciprocity of relative velocities, they understand (p. 124):

... that the velocity of K with respect to K' be equal to minus the velocity of K' with respect to K for all pairs of inertial systems ...

(ii) Their claim (p. 125, n. 8) that:

... by McPhee's Theorem, different values of ϵ in different inertial systems entail different values for the respective relative velocities.

(iii) Their supposition that *non*-reciprocity of the relative velocities violates the strong principle of relativity.

If the arguments which E & B base on either version of the principle of relativity are not to beg the question vis-à-vis my analogy, they must succeed in showing the following: the latitude for conventional choices of nonstandard synchronisms ϵ within the range $0 < \epsilon < 1$ is circumscribed by those components of the principle of relativity which are clearly factual and do not already implicitly contain particular conventions. For it would clearly beg the question to indict my statement by doing the following: pointing to the incompatibility of my statement with versions of the principle of relativity that already embody some *one* convention to the exclusion of all the others which I countenanced as alternatives. That either version of the principle of relativity contains conventional components in addition to factual ones could hardly be denied by anyone. Two examples will illustrate the distinction between factual and conventional ingredients of the principle of relativity.

(1) Let us note one use which Einstein makes of the principle in §2 of his basic 1905 paper on the STR. There he invokes the principle of relativity to infer the result of only the *first* of two different operations of measuring length. And he mentions the result of a second operation which will enable us to clarify a misunderstanding of the principle of relativity below. He writes:

Let there be given a stationary rigid rod; and let its length be l as measured by a measuring-rod which is also stationary. We now imagine the axis of the rod lying along the axis of x of the stationary system of coordinates, and that a uniform motion of parallel translation with velocity v along the axis of x in the direction of increasing x is then imparted to the rod. We now inquire as to the length of the moving rod, and imagine its length to be ascertained by the following two operations:

(a) The observer moves together with the given measuring-rod and the rod to be measured, and measures the length of the rod directly by superposing the measuring-rod, in just the same way as if all three were at rest.

(b) By means of stationary clocks set up in the stationary system and synchronizing in accordance with §1 [$\epsilon = \frac{1}{2}$], the observer ascertains at which points of the stationary system the two ends of the rod to be measured are located at a definite time. The distance between these two points, measured by the measuring-rod already employed, which in this case is at rest, is also a length which may be designated "the length of the rod."

In accordance with the principle of relativity the length to be discovered by the operation (a)—we will call it "the length of the rod in the moving system"—must be equal to the length l of the stationary rod.

The length to be discovered by the operation (b) we will call "the length of the (moving) rod in the stationary system." This we shall determine on the basis of our two principles, and we shall find that it differs from l ([3], pp. 41–42).

In this particular context, the factual content of the principle of relativity (strong version) is that the *number of coincidences* of the measuring rod with the rod to be measured is *invariant* with respect to different inertial systems S and S' . But this factual state of affairs will issue in the same numerical length l in S and S' , as here required by Einstein, only if we adopt the *convention* of letting the measuring rod be unity (or the same multiple of unity) in *both* S and S' . And if someone were to change the unit of length in going from S to S' and thereby obtain correspondingly different lengths in S and S' from the measurements, what would be the force of indicting these findings as incompatible with the principle of relativity? It would demonstrate merely that the conventional choice of different units of length in S and S' will not yield the same description of the given facts of rod coincidences as the choice of the same unit in both frames, since these alternative choices embody incompatible conventions.

(2) Consider the weak version of the principle of relativity which demands the *formal invariance* of the laws of nature with respect to inertial systems as linearly *related by the Lorentz transformations*. *Would this principle be violated by using polar space coordinates ρ and θ in one or more of the inertial systems while using rectangular coordinates x , y , and z in the others? Using rectangular space coordinates in one frame, the law of light propagation in that frame is*

$$x^2 + y^2 + z^2 - c^2 t^2 = 0,$$

whereas in another frame using polar space coordinates and a time coordinate t_0 , it has the *different form*

$$\rho^2 - c^2 t_0^2 = 0.$$

Clearly, the space and time coordinates of the two frames, via which these two formally different equations would be intertransformable, would *not* be related by the linear Lorentz transformations. But presumably no one would say that such a state of affairs is incompatible with the *factual physical content* of the given version of the principle of relativity.

The significance of our second example is as follows: pending *other physical arguments* against $\epsilon \neq 1/2$, E & B 's appeal to its violation of the *weak* (Einstein) principle of relativity does *not* establish the incompatibility of $\epsilon \neq 1/2$ with any *factual physical ingredients* of the weak version. Indeed, the objective physical facts asserted by the STR do not make it incumbent on any exponent of that theory to espouse the weak version of its First Postulate, unless he also adopts certain conventions.

The difficulties with E & B 's invocation of the *strong* (Poincaré) version are more ramified and call for three sets of comments as follows:

1. As it was formulated by E & B , the strong version of the principle of relativity is false. For it avowedly "rules out differences in quantitative determinations (say, of . . . velocity . . .) between different inertial systems" (p. 122, n. 6), presumably in regard to ingredients of the laws of nature. Now, while this principle holds for the velocity of propagation of any one light wave *in vacuo*, it does *not* also hold for the velocity of any one sound wave in air (under standard conditions of pressure

and temperature). To see this, we must clearly distinguish the following two very different kinds of sets of determinations of the velocity of sound, which are the respective velocity counterparts of Einstein's different length measuring operations (b) and (a) above:

Group α of Measurements. A sound wave is emitted in the air of a frame S under conditions of pressure and temperature which are standard *as measured in S* . By "the air of a frame S ," we mean the air which is macroscopically at rest in S . The S observer finds that under these conditions, the velocity of his sound wave in S is approximately $\frac{1}{3}$ of a kilometer per second. Let an observer in the different inertial system S' determine the velocity of *that same sound wave*, along with determining the prevailing relevant conditions of pressure and temperature. Clearly, the velocity of that same sound wave as measured in S' will *not* be the same numerically as in S , any more than Einstein's "length of the (moving) rod in the stationary system" has the same numerical value l as his "length of the rod in the moving system." Furthermore, whereas the aerostatic pressure as measured by S' is the same as the pressure measured in S , at least one authority on the theory of relativity, R. C. Tolman, has argued that the temperature is not a numerical invariant ([17], pp. 154, and 158–159).

On the other hand, if the velocity of some one particular light wave were analogously measured in both S and S' , it would, of course, be found to be numerically the same. Thus, as Tolman puts it:

The expression of the equations of physics in a *form* which is independent of the co-ordinate system does not in general prevent a change in their numerical content when we change from one system of coordinates to another ([17], p. 166).

And this result demonstrates the falsity of the strong principle of relativity *as formulated by E & B*.

Group β of Measurements. A sound wave is emitted in the air of a frame S under standard conditions of temperature and pressure as measured in S . And *another* sound wave is also emitted in the air of a different frame S' under the same standard conditions as measured in S . Assume that the principle of relativity asserts the nonexistence (and hence undetectability) of the absolute uniform translatory motion of an inertial system. Then, *if S and S' each use standard synchronism*, this principle guarantees that S and S' will each obtain the same numerical value for the velocity of sound, provided that they also each use the same units of length and duration.

2. *E & B* confuse the *intrasystemic* principle of relativity with an *intersystemic* condition of reciprocity when demanding the reciprocity of relative velocities on the strength of the strong principle of relativity. The distinction between these two principles has been lucidly explained by H. Mehlberg as follows:

Any principle of relativity states that the intrasystemic laws of nature are the same in a specifiable class of frames of reference, say, in inertial frames. A principle of reciprocity, implicit in several discussions of relativity although hardly ever formulated with any precision, requires that whenever a law of nature asserts a relation R between the frame K and the frame K' (both members of a specifiable class of frames of reference) there must be a companion law to the effect that the very same relation R obtains also between K' and K . Thus, according to the special theory of relativity, if a physical object is a sphere in one inertial frame K , it will in general be an ellipsoid in any other frame K' . The principle of

reciprocity states then, in addition, that a physical object whose shape is a sphere in K' will have the shape of an ellipsoid in K

It goes without saying that the principle of reciprocity does not coincide with the principle of relativity since the latter applies only to intrasystemic laws whereas the former is primarily relevant to intersystemic laws.

. . . To put it otherwise, it is a fundamental claim of special relativity that all intrasystemic laws of nature are the same in an ostensively defined equivalence class of frames of reference which move relatively to each other in straight lines with uniform speed. This postulate, however, must not be confused with the stronger requirement of the Lorentz invariance of all laws of nature in coordinate systems associated with every equivalence class of admissible frames of reference. . . . It remains to be seen whether intersystemic laws can bridge the gap between the major postulate of special relativity and the stronger requirement of Lorentz invariance ([12], pp. 473 and 474).

Let us grant the truth of the factual components of the principle of reciprocity as stated by Mehlberg no less than of those of the principle of relativity. Note incidentally that E & B 's condition of reciprocity of velocities differs from Mehlberg's principle of reciprocity. A *law of nature* assures the existence of the relation R whose reciprocity is asserted in Mehlberg's formulation. But no law of nature asserts that the numerical value of the velocity of S' with respect to S is v . Hence the relation R for which E & B 's condition claims reciprocity obtains merely *de facto*.

Nonetheless, E & B 's reciprocity of relative velocities is a *non-trivial* consequence of the Lorentz transformations, which presuppose $\epsilon = 1/2$.³ But what follows from the fact that E & B 's reciprocity of velocities is thus a necessary condition for the validity of the Lorentz transformations? Does this show that a violation of velocity reciprocity also violates the *factual* physical commitments ingredient in the principle of relativity as augmented by the factual content of the Lorentz transformations? That the answer is negative becomes evident from the fact that there are violations of velocity reciprocity which incontestably arise solely from the use of alternative conventions. For suppose that it *were* granted that the assertion of the simultaneity corresponding to $\epsilon = 1/2$ is a factual truth rather than a matter of convention and that each of the frames S and S' employs the standard synchronism. But suppose also that S employs the English system of units for length while S' uses the metric system. Even if S and S' proceed alike to measure velocity in all *other* respects (e.g. they choose the same unit of time), a violation of numerical velocity reciprocity will result!

3. It will be recalled why E & B were concerned to claim that the strong principle of relativity requires reciprocity of relative velocities. They did so with a view to inferring that nonstandard synchronisms $\epsilon \neq 1/2$ violate the strong principle of relativity because such synchronisms allegedly rule out velocity reciprocity. We must therefore still examine their claim that (p. 125, n. 8):

. . . by McPhee's Theorem, different values of ϵ in different inertial systems entail different values for the respective relative velocities.

By E & B 's own account of McPhee's result, the conclusion that clocks must be in standard $\epsilon = 1/2$ synchrony follows from a conjunction. *One* of the pertinent con-

³ For a statement of the relevant derivation, see [1]. Bergmann points out that the consequence is nontrivial because "neither the unit of length nor the unit of time is directly comparable" in S and S' .

juncts is the condition of velocity reciprocity. The other conjunct is E & B 's requirement that any motion which is both uniform and rectilinear in one frame S will likewise be so in any other frame S' . They call this requirement the "condition of linearity" (p. 123). In other words, this condition demands the *invariance* of Newton's first law of motion with respect to inertial systems. It is plain that McPhee's result does not entail that nonstandard synchronisms must violate velocity reciprocity. What it does entail is the weaker conclusion that $\epsilon \neq \frac{1}{2}$ violates *either* the condition of linearity or the condition of velocity reciprocity. And E & B are not entitled in this context simply to take the condition of linearity for granted without regard to whether the assertion of the invariance of Newton's first law embodies conventional as well as factual ingredients.

To ignore the latter issue by taking this invariance for granted would beg the question here. For E & B are endeavoring to rule out (or at least circumscribe) non-standard synchronisms by adducing the factual components of physical principles to the exclusion of conventions implicit in these principles. In particular, they invoked McPhee's result in an attempt to refute my statement that "the use of different values of ϵ in different inertial systems S and S' can no more violate the First Postulate of the STR than the use of centimeter units to express the velocity of light in S and of miles to express that same velocity in S' ." But I had shown in detail that the invariance of Newton's second and first laws of motion depends not only on facts but also, among other things, on the particular *time-congruence* that is used to render them ([9], Ch. 2, pp. 66–73, esp. pp. 71–72). And furthermore I had argued that assertions of congruence among intervals of space and time importantly contain a conventional ingredient ([9], Ch. 1; also [10], §§2 and 6). E & B offer an argument against this thesis concerning congruence in connection with their contention that slow clock transport establishes the nonconventionality of the standard synchrony $\epsilon = \frac{1}{2}$. Our section 3.2 below will give reasons, however, for rejecting this argument of theirs. Hence we can deny that E & B are entitled to take their condition of linearity for granted in this context.

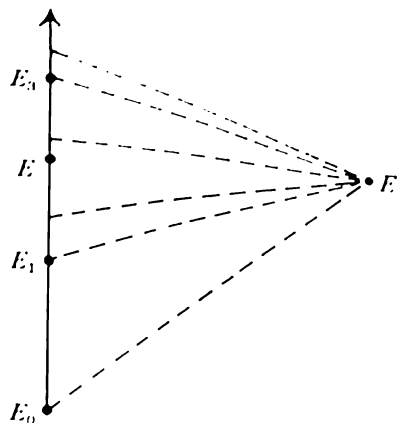
But we saw that without the assumption of the latter condition, McPhee's result does *not* yield E & B 's conclusion that nonstandard synchronisms must violate velocity reciprocity. And we also saw above that they likewise failed to show that the factual content of the strong relativity principle requires velocity reciprocity. Yet these two unfounded conclusions were the basis for their contention that non-standard synchronisms violate the strong principle of relativity and that my analogy is "in error" (p. 122, n. 5).

3. The Philosophical Status of Simultaneity by Slow Clock Transport in the STR.

The physical basis of E & B 's main thesis is the following: in the STR, the behavior of clocks under transport assures that clocks in slow clock transport synchrony are automatically in standard signal synchrony $\epsilon = \frac{1}{2}$. And their main thesis itself is the philosophical claim that in the STR the relations of simultaneity corresponding to $\epsilon = \frac{1}{2}$ obtain as a matter of temporal fact. Therefore, E & B reject as false the view which Reichenbach had espoused and attributed to Einstein. For on that view, relations of distant simultaneity in the inertial frames of the STR involve a con-

ventional ingredient in an important sense. In opposition to this view, *E* & *B* claim that the simultaneity relations resulting from slow transport synchrony cannot involve a conventional ingredient in "any but a most trivial sense" (p. 127).

In order to articulate the philosophical status of the simultaneity relations corresponding to slow clock transport synchrony in the inertial frames of the STR, we shall first characterize the status of simultaneity in two other universes as follows: (1) the world of Newtonian mechanics and (2) a fictitious universe to which we shall refer as "quasi-Newtonian." The latter resembles Newton's world in that transported clocks have the same readings whenever they meet again after being *locally* synchronized with one another and then transported arbitrarily to the same place elsewhere. But the second universe differs from Newton's world in that light *in vacuo* and gravitational influences are the fastest causal chains which can link any two space points in any inertial frame of the quasi-Newtonian universe, and each



of these two influences has the same *finite* round-trip velocity in any one inertial frame.

1. *Simultaneity in Newtonian Mechanics.* Let the solid line on the left be a portion of the world-line of a clock U_1 which is at rest at a point A of an inertial system I . And let E' be an event belonging to the career of another clock U_2 at rest at a point B of I . Furthermore, suppose that any clock U which moves in I and intersects the world-line of U_1 has the same reading as the latter for the event of their first encounter. It is then a fact that (after allowance for the effects of what Reichenbach has called "differential forces") U will have the *same reading* as U_1 for any *subsequent* encounter. This *agreement* between U and U_1 is not, however, the sole respect in which Newtonian and relativistic clock transport *differ* from one another.

In the Newtonian world of arbitrarily fast particles (or causal chains), the career S of U_1 contains a *unique* event E which cannot *also* belong to the career of any moving clock U (or other particle) containing E' . Once the Newtonian time system is elaborated, this fact can be expressed by the statement that "the same body (U) is elaborated, this fact can be expressed by the statement that "the same body (U) cannot be at two different places (A and B) at the same time." And the specified unique event E divides S into disjoint open subintervals of events X and Y having the following properties: *every* event x in X and *every* event y in Y can also belong

to the world-line of a moving clock U whose intersection with the world-line of U_2 is E' . Furthermore, if each clock U was *locally* synchronized with U_1 , the time t' of E' on *every* U is the same and is numerically *between* the time of x on U (or U_1) and the time of y on U (or U_1). The world-lines of such clocks U are shown by dotted lines. And the betweenness of E' on these world-lines is a matter of purely ordinal temporal fact. For it does *not* depend on invoking any durational measure of an event interval xE' or $E'y$. Thus, for *any* x in X and *any* y in Y , E' is temporally between them on the basis of the identical reading t' of suitably fast moving clocks U whose respective careers likewise comprise x and y . And E' is the only event on the world-line of U_2 sustaining these betweenness relations to *all* of the members of X and Y .

But is also true (by our definition of X and Y) that E is the only event in S which is temporally between every x in X and every y in Y . It follows that (i) E' and E are temporally between identically the same events in S , and (ii) in any system of quasi-serial temporal order comprising the events on U_2 and in S , E' and E occupy the same place with respect to the order of earlier and later *as a matter of ordinal temporal fact*. Hence on the basis of temporal betweenness relations alone, E' is uniquely simultaneous with E within S , and E is uniquely simultaneous with E' within the career of U_2 .

Note that these relations of temporal betweenness and *simultaneity* can be asserted without any prior appeal to a time metric that assigns durational measures to pairs of events or event-intervals: the ascription of simultaneity to E and E' on the basis of the time numbers furnished by the clocks U_1 and U did *not* require the *prior assumption* that E' is later than some event E_0 in S by the same amount as E . And thus this assertion of simultaneity did not first need to *assume* that the *duration measures* of the event intervals E_0E and E_0E' are equal. Instead, it is a fact that E and E' are simultaneous on purely ordinal grounds. This fact renders the event intervals E_0E and E_0E' doubly coterminous in the *time* continuum. For it assures that E and E' belong to the same instant in the time continuum. And the double coterminousness of E_0E and E_0E' in the *time* continuum guarantees the equality of their durations, since these are each the measure of the self same interval in the time continuum.

By the same token, if E'_0 belongs to the world-line of U_2 and is simultaneous with E_0 on the basis of our purely ordinal clock transport criterion, then the durational measures of the event intervals E_0E and E'_0E' will be equal simply because they are each the measure of one and the same interval of instants in the time continuum. But it is clear that the ordinal facts of the temporal order which make for these particular durational equalities do *not* entail congruences among successive time intervals on any one world-line or in the time continuum at large. Hence the particular durational equalities which are deducible from the simultaneity relations in Newton's theory do not serve to impugn the thesis that the intervals of the time continuum are devoid of intrinsic congruences.⁴ And there is no ordinal or topological basis that would yield such congruences.

It will turn out to be very important for our assessment of E & B 's claims to note

⁴ For a detailed statement of that thesis, see [7], Ch. III, §§2.9 and 2.10.

at this point how much is required to assure the non-metrical, purely ordinal character of the unique simultaneity relation between E and E' as furnished by Newtonian clock transport. Specifically, suppose that, *contrary* to Newtonian physics, there is a positive closed subinterval E_1E_2 of S , however small, which contains E but whose members x and y *cannot* be on the world-lines of any clocks U that contain E' . And suppose further that, as in Newton's physics, all clocks U capable of linking *other* events in S to E' by transport yield the *same* time t' on coinciding with E' , where t' is equal to the time t read by the clock U_1 upon the occurrence of E . Then the mere agreement among such clocks U in regard to reading the time coordinate t' , where $t' = t$, would *not* suffice to establish the unique simultaneity of E' with E on purely *ordinal* grounds. For under the now posited hypothetical circumstances, there is an ordinal hiatus. And therefore this assertion of simultaneity would *also* depend crucially on the *metrical* claim that E and E' are *durationally equidistant* from every given event in S that is earlier than E_1 or later than E_2 . Without such a metrical claim, there is now *no* ordinal basis for asserting that E' sustains any relations of temporal betweenness to pairs of events belonging to the closed interval E_1E_2 . And hence there is then no purely ordinal basis for claiming that E and E' are temporally between identically the same events in S . The mere numerical equality of the clock numbers t and t' cannot remedy this ordinal hiatus and thus cannot render E and E' simultaneous without appealing to durational equidistance.

In order to characterize further the philosophical status of the simultaneity furnished by Newtonian clock transport, it behooves us to comment on the bearing of causal relations in Newton's theory on its time relations. Newton's third law of motion (law of action and equal, opposite reaction), coupled with his law of universal gravitation tells us that our E and E' are linkable by reciprocal instantaneous gravitational influences. These can be represented as causal chains $EE'E$ or $E'E'E'$ whose "emission" at either E or E' coincides with their "return" to either E or E' . And since no Newtonian body can be at two different places simultaneously, no Newtonian body or clock can link E to E' so that these events coincide spatio-temporally with event-members belonging to its career. Indeed, in Newton's world, gravitational influence chains are the only causal chains whose careers can include simultaneous events such as E and E' . And gravitational chains comprise none but simultaneous events. Moreover, any set of pairwise *non-simultaneous* events can be linked by a *non-gravitational* causal chain which is genidentical, i.e. which is constituted by the career of one and the same body. The career of a single standard clock is, of course, an instance of merely one particular species of genidentical causal chain.

It would clearly be inconsistent with Newton's temporal order to demand, as is done in the STR, the *non-simultaneity* of two events connectible only by the fastest causal chain rather than by a single clock. For on Newton's theory, our events E and E' are simultaneous according to its clock readings, and yet they are connectible by Newton's fastest causal chain (gravitation) and only by such a chain. By contrast, the STR requires its clocks to be set so as to issue in the *non-simultaneity* of any two events which can belong only to the career of its fastest causal chains (light), even though these events cannot both be on the world-line of a single clock.

Thus, in the STR, *any* two causally connectible events whatever are to be assigned different time numbers by clocks in every inertial system. This requirement of the STR is its coordinative definition for the invariant *time-separation* or *non-simultaneity* of two events, and will be called “N-S.” Since this coordinative definition N-S is inconsistent with Newton’s theory, we must not allow ingrained relativistic habits to introduce it tacitly, when now proceeding to consider the bearing of causal relations on time relations in Newton’s world.

A word of caution is needed concerning my characterization of N-S as the STR’s coordinative definition of *invariant* time-separation of two events. This has meant traditionally that the particular time order of two causally connectible events is *invariant* with respect to all inertial frames. Recently, however, there have been articles asserting the physical possibility of causal chains faster than light *in vacuo* linking events whose time-order is *not* invariant according to the Lorentz transformations of the STR [cf. G. Feinberg, “Possibility of Faster-Than-Light Particles,” *The Physical Review*, vol. 159, 1967, pp. 1089–1105; R. G. Newton, “Causality Effects of Particles That Travel Faster Than Light,” *Physical Review*, vol. 162, 1967, p. 1274; S. A. Bludman and M. A. Ruderman, “Possibility of the Speed of Sound Exceeding the Speed of Light in Ultradense Matter,” *Physical Review*, vol. 170, 1968, pp. 1176–1184, and M. Ruderman, “Causes of Sound Faster Than Light in Classical Models of Ultradense Matter,” *Physical Review*, vol. 172, 1968, pp. 1286–1290]. These publications pose issues that cannot be treated here. Hence in this paper, I shall take no cognizance of such modifications of N-S or even of the STR as may be required by these results.

See Append
§17 for
comments
on super-
light causal
chains

I have shown elsewhere in detail that certain relations of *causal betweenness* are exhibited by the events belonging to genidentical chains which are *not* spatially self-intersecting in at least one inertial frame ([8], pp. 56–62). Such relations of causal betweenness obtain alike in the universes of Newtonian mechanics and of the STR.⁵ But in Newton’s world the events linked by his *gravitational* chains do not exhibit these relations of causal betweenness. The latter chains are spatially self-intersecting *in every frame*, since the emission of any Newtonian gravitational influence coincides (spatio-temporally) with its return in every frame. Therefore each gravitational chain is spatially self-intersecting in any frame whatever. Now, it is possible to characterize the relations of *temporal* betweenness and simultaneity in Newton’s world on the basis of the causal betweenness defined by the following subclass *K* of its physically possible causal chains: *K* is the set of all those genidentical chains which are not spatially self-intersecting in at least one inertial frame. Although our clocks *U* all have world-lines which *are* spatially self-intersecting at the space point *A* in frame *I*, the careers of these clocks *U* *do* qualify as members of *K*, since they are not spatially self-intersecting in inertial frames which move in a direction perpendicular to the spatial straight line *AB* in *I*.

Being mindful of the important fact that *K* also has members other than the careers of standard clocks, we can state the relevance of the entire membership of

⁵ The relations of causal betweenness in question do not entail but merely allow the introduction of the STR coordinative definition N-S. And they do not, of course, depend on N-S as a premiss.

K to Newtonian temporal relations amongst events as follows: the relations of *causal* betweenness furnished by the members of K are completely isomorphic with the relations of *temporal* betweenness furnished by the time numbers on Newtonian clocks, whose careers are likewise members of K . And the thus K -defined relations of temporal betweenness also permit us to say that our events E and E' are temporally between identically the same events in S . Therefore, E and E' also turn out to be ordinarily simultaneous on the more general basis of the causal relations furnished by the membership of K , which is far wider than that of the class of the careers of clocks. Indeed, we shall see at the end of this paper that the *numerical* betweenness of the coordinates furnished by a clock derives its *temporal* significance from the underlying causal betweenness. Although a Newtonian gravitational chain may well qualify as being genidentical, it does not qualify as a member of K . Hence the K -defined relations of temporal betweenness do not allow the deduction of the paradoxical conclusion that either of two simultaneous events belonging to a gravitational chain is temporally between the other and some third event. Nor is the STR coordinative definition N-S available in Newton's theory to permit the inference that E and E' are paradoxically nonsimultaneous.

It is clear from our analysis that in Newton's world events are simultaneous as a matter of physical fact because of nonmetrical relations of temporal betweenness furnished by that world's clocks and/or causal relations in K . Spatially separated Newtonian clocks at A and B can be consistently synchronized by transporting a third clock U from A to B and making each of them locally synchronous with U when it coincides with them. We see that the sameness of the time numbers furnished for simultaneous events by such synchronized clocks A and B renders an equality relation that exists between these events as a matter of physical fact. Thus the existence of the relation to which the Newtonian theory applies the name "simultaneous" does not involve any conventional ingredient. What is conventional here is the particular identity of the time number assigned alike to all members of a class of simultaneous events. The identity of that number results from one arbitrary setting of one clock. But the equality relation of simultaneity rendered by the same clock numbers is not predicated on a convention in Newton's theory.

Newtonian simultaneity is absolute in the standard physical sense that the simultaneity of two events E and E' is *invariant* with respect to all reference frames. But Newton's simultaneity is also factual, as opposed to conventional, because it is vouchsafed by purely ordinal temporal facts. We shall soon see apropos of the fictitious *quasi*-Newtonian world that although *its* simultaneity relations are also *invariant*, they do depend on a metrical appeal to the durational equidistance of E and E' from one or more other events. And we shall maintain against E & B that this dependence on a time metric introduces a nontrivial conventional ingredient into the simultaneity relations of the quasi-Newtonian world. By contrast, the simultaneity relations of Newton's world are nonconventional.

It is evident that, from the foundational point of view, a moving entity has a determinate *one-way* velocity in a given frame of Newton's theory only because a numerical one-way transit time is already defined in the theory. And if that one-way transit time is to have any significance for the purposes of physical theory, it must

be furnished by *synchronized* clocks. Synchronized clocks, in turn, are those that read the same time numbers for events which qualify as simultaneous within the given physical theory. Hence in Newton's theory no less than in the STR, a one-way velocity presupposes relations of distant simultaneity. But it is clear from our analysis that the existence of Newtonian simultaneity relations is not similarly predicated on any numerical one-way velocities. Thus Newton's simultaneity does not pose a problem of avoiding a vicious logical circle.⁶ Of course, for epistemic purposes the simultaneity of two events may be *inferred* in some circumstances in Newton's theory from information containing the value of a one-way velocity.

2. *Simultaneity in the Quasi-Newtonian Universe.* The quasi-Newtonian world differs from Newton's in the following respect: in the former, light *in vacuo* and gravitational influences are the fastest causal chains which can link any two space points in any inertial frame, and each of these influences has the same *finite* round-trip velocity in any one inertial frame. Let the dotted lines E_1E' and $E'E_3$ in our diagram be the world-lines of light rays. Then in the quasi-Newtonian world, none of the events in the *open* interval E_1E_3 are connectible to E' by *any* causal chain, and none of the events belonging to the *closed* subinterval E_1E_3 of S are connectible to E' by clock transport. But nonetheless in this world, it is a fact that all transported clocks U capable of linking *other* events in S to E' spontaneously *agree* in yielding the same time number t' upon coinciding with E' . Assume that this number t' is the same as the number t read by the clock U_1 upon the occurrence of E , a sameness which is invariant in all frames of this fictitious world, and does not depend on the specification of any one-way velocity of a clock. Pending the introduction of the *metrical* relation of *durational equidistance* further on, the temporal relations yielded by the quasi-Newtonian facts asserted thus far are of a purely *ordinal* kind. And since these ordinal relations obtain on the strength of relations of causal betweenness, they hold as a matter of objective physical fact. Further on, when the metrical relation of durational equidistance is made available by explicit introduction, this relation will be seen to obtain as a matter of convention and *not* as a matter of physical fact.

It is highly important to see now that a further assumption is required in this context to permit the inference that E' must be simultaneous with E , i.e. that E' must occupy the same place as E in the system of temporal order defined by the career of the clock U_1 . Each clock U was (and will be) *locally* synchronized with U_1 at the intersection of their world-lines by reading the same time number upon their encounter. The clocks U as well as the clock U_1 each exhibit the serial temporal order of the events *on their own respective world-lines*. And the light ray $E_1E'E_3$ defines relations of temporal betweenness on its own world line. In particular, E and E' are both between E_1 and E_3 such that $t = t'$. But in crucial contrast to the Newtonian world, there is *no physical basis* here for asserting that E' is also temporally between pairs of events x and y such that x belongs to the *open* interval E_1E , and y to the open interval EE_3 . And, therefore, there is the following contrast with Newton's world in regard to the deducibility of the simultaneity of E and E' : the purely *ordinal* significance of the equal coordinates t and t' , which U_1 and U assign

⁶ Reichenbach was clear on this point: see his discussion on pp. 129 and 146–147 of [14].

respectively *on their own world-lines* to the spatially separated events E and E' , is *not* sufficient at all to order E and E' in relation to each other as uniquely belonging to the same instant within a *common* system of quasi-serial temporal order!

In order to make the latter assertion of simultaneity for the quasi-Newtonian world, we must couple the statement of the *local* synchronism of U_1 and U with a *metrical* claim as follows: E and E' are *durationally equidistant* from one or more other events such as E_0 , which is at the intersection of the world-lines of U_1 and of one of the clocks U at which these clocks were locally synchronized. For we just saw that E' is ordinarily indeterminate with respect to E , despite the identity of their time coordinates. And that indeterminateness is supplanted by simultaneity here only by adding that since $t = t'$ and $t - t_0 = t' - t_0$, the *same amounts of time* have elapsed on the world-lines of U_1 and U in the course of their spatial separation after they each read t_0 at their intersection E_0 . Hence a *durational measure* is being invoked according to which equal differences in the time coordinates of two event pairs assure the equality of their respective durational measures. And since the time interval measures of E_0E and E_0E' are then equal along the two world-lines, E and E' are *durationally equidistant* from E_0 . Thus the identity of the time coordinates t and t' serves to establish the simultaneity of E and E' here only via a metrical appeal to the *durational congruence* of intervals on different world-lines. But as we shall now see, precisely this dependence on a time *metric* nontrivially imports a *conventional* ingredient into the simultaneity relation of E to E' .

Like points of physical space, instants of time are unextended and lack inherent magnitude. The career of the center of mass of the clock U generates a continuum of instants, and the clock as a whole assigns ordinal time labels or coordinates to these instants. In his Inaugural Dissertation, Riemann called attention to the following important fact pertaining to such an interval: the continua constituting various intervals of instants of time or points of space have no intrinsic properties which could serve as a basis for a built-in measure of their respective extensions. Thus, the intrinsic property of cardinality cannot furnish such a basis, since all (nondegenerate) intervals have the same cardinality. And the intervals of instants corresponding to our event pairs E_0E and E_0E' have no intrinsic durational measures which could attest to their equality.

Hence there is nothing in the intrinsic make-up of these intervals that could certify it to be a matter of temporal fact that the clocks U_1 and U are *metrically isochronous* or have equal "rates" for segments of their world-lines which do not coincide. In this important sense, the isochronism of the clocks on which the durational equidistance of E and E' from E_0 is predicated involves a convention. Therefore, in contrast to Newton's world, the assertion that E and E' are simultaneous in the quasi-Newtonian world rests on a convention in a *non-trivial* sense. And only the identity of the particular time number assigned alike to both E and E' is *trivially* a matter of convention, just as in the Newtonian world.

Thus Riemann's analysis of the status of congruence among intervals of points or instants enables us to see why the assertion of temporal equidistance here rests on a convention. And hence our account vindicates Reichenbach's claim that the simultaneity of E and E' in the quasi-Newtonian world is a matter of "definition"

The
impending
statement
of conven-
tionality is
supplanted
by ch. 16,
Section 3 (ii).
See Appendix
§ 53

(as opposed to “fact”), even though it is a *fact* that all clocks U agree in reading the same number t' upon the occurrence of E' ([14], pp. 133–135). He adds usefully that we are dealing here with “a *definition* of simultaneity, which is a definition in the same sense as the definition of congruence by means of rods” ([14], p. 135). And we can say further that relatively to the adoption of the simultaneity convention based on clock transport, in conjunction with a spatial metric, the one-way velocities of light become matters of empirical fact. This is altogether analogous to saying that relatively to the convention that a given rod is self-congruent under transport and is a unit rod, the lengths of other bodies become matters of empirical fact.

E & B have challenged this conclusion in the sense of indicting this sense of conventionality of simultaneity as “trivial” and hence “as not worth discussing” (p. 135). And their basis for this indictment is as follows (p. 126):

Reichenbach claims . . . that even if the predictions of the Special Theory of Relativity were wrong . . . and it were possible to establish a synchrony by a transport procedure, it would still be a matter of definition, and not of empirical fact, that clocks that are locally synchronous remain synchronous after separation. For the case would then be entirely analogous to that of distant congruence. . . . we could still maintain that locally synchronous clocks do not remain synchronous after separation, even if they would be found to be synchronous whenever and wherever they are brought together again. . . . But in this way every relationship of quantitative equality that depends upon local comparison is conventional. Distant mass or temperature equality would be no less conventional in this sense than distant simultaneity.

And when speaking of the simultaneity relations furnished by *slow* clock transport in an inertial system of the STR, E & B assert (p. 127):

. . . if the empirical predictions of the Special Theory of Relativity regarding clock transport should prove to be correct, then there is a physical relationship that is in fact symmetrical and transitive and which could be used to define distant simultaneity. This physical relationship is independent of any signaling procedure, and does not require for its determination any prior measurements of velocity. It is therefore like the relationships that may be used to determine distant mass equality, which is not held by anyone to be conventional in any but a most trivial sense.

(p. 134):

While it may be agreed that slow transport synchronization is possible in any inertial system, it may be argued that we have no more reason to accept the slow transport definition of synchrony that we have given than to accept any nonstandard slow transport definition. Thus we might allow that infinitely slowly transported clocks, once synchronized, will always be found to be synchronous with each other, but we might deny that they remain synchronous while they are separated. But this only shows that distant simultaneity is conventional in the trivial sense that any quantitative equality between two things at a distance is conventional. If this is all there were to Reichenbach’s conventionality thesis, it would be absurd to devote so much time to discussing it.

These remarks by E & B call for critical comment. It is *not* true that in the *quasi*-Newtonian universe, simultaneity is conventional *only* in the way in which “every relationship of quantitative equality that depends upon local comparison is conventional” (p. 126). We saw that in the Newtonian universe, the simultaneity relations furnished by transported clocks hold as a matter of ordinal temporal fact. But these relations depend on the existence of intersections of the world lines of U_1

and U at which these clocks are *locally compared* or synchronized. Though clearly depending on local comparisons of clocks, Newton's simultaneity relations are *not* conventional. While these Newtonian relations are not themselves even trivially conventional, the particular identity of the numerical *names* which happen to be used to convey these simultaneity relations is, of course, so. But the trivial conventionality of the particular time coordinate (numerical name) assigned alike to two or more events belonging to a particular simultaneity class does not at all render the simultaneity relation among these events itself conventional. And the identity of that particular time coordinate (numerical name) can be determined by the choice of a zero of time on one clock or of its equivalent. By contrast, we saw on the basis of Riemann's philosophy of metrical congruence, that in the quasi-Newtonian universe, the simultaneity of E and E' rests on a nontrivial convention, over and above the trivial one involved in the choice of a zero of time.

The conventional element that is common to simultaneity in both the Newtonian and quasi-Newtonian worlds is a zero of time or its equivalent on one clock and is indeed only trivial. But it is fallacious to infer with E & B that simultaneity in the quasi-Newtonian world and simultaneity by slow clock transport in the STR involve none but this trivial conventionality. E & B commit this fallacious inference, because they mistakenly assume that the dependence on local numerical comparison must be the *sole* source of conventionality. It is this erroneous assumption which prompts their attempt to trivialize the conventionality of non-Newtonian simultaneity by invoking the irrelevancy that "Distant mass or temperature equality would be no less conventional in this sense than distant simultaneity" (p. 126).

Besides, even if it *were* true that "every relationship of quantitative equality that depends upon local comparison is conventional" *in one and the same sense*, it would not follow that the *generality* of this conventionality establishes its triviality. It does not trivialize the attribute of mortality to point out that all men possess it alike.

Our comparison of the Newtonian and quasi-Newtonian worlds has shown that absolute simultaneity relations which are furnished by clocks *can* be nontrivially conventional. And this fact prompts me to correct an earlier statement of mine in which I claimed that absolute relations of simultaneity *must* be nonconventional. In a paper of 1961, I properly characterized Einstein's conception of simultaneity in the STR as resting on two assumptions: (i) within the class of physical events, standard clocks do *not* define relations of absolute simultaneity under transport, since clocks which are initially synchronized at one place A will *generally* no longer agree after being transported to another place B , and (ii) light is the fastest signal *in vacuo* in the following ordinal sense: no kind of causal chain (moving particles, radiation) emitted *in vacuo* at a given point A together with a light pulse can reach any other point B earlier—as judged by a local clock at B which merely orders events there in a metrically arbitrary fashion—than this light pulse.

Commenting on these two assumptions, I first said correctly:

... if assumption (i) had been thought to be *false*, then the belief in the truth of (ii) would *not* have warranted the abandonment of the received Newtonian doctrine of absolute simultaneity ([6], p. 45).

But then I went on to infer unsoundly:

And, in that eventuality, the members of the scientific community to whom Einstein addressed his paper of 1905 would have been fully entitled to *reject* his *conventionalist* conception of *one-way* transit times and velocities.⁷

But absoluteness of simultaneity in the physical sense relevant here means intra- and intersystemic concordance in the simultaneity verdicts of all transported clocks. And we saw apropos of the quasi-Newtonian world that such absoluteness of simultaneity is not sufficient for its nonconventionality.

My statement of the status of simultaneity in the quasi-Newtonian world now enables us to assess simultaneity by slow transport synchrony in the STR.

3. *Simultaneity by Slow Clock Transport Synchrony in the STR.* The universe of the STR differs from the quasi-Newtonian world in the following two respects: (1) clocks which are in slow transport synchrony in the various inertial systems of the STR yield the time coordinates of the Lorentz transformations. Hence the simultaneity verdicts of such clocks are *intersystemically* discordant or relative rather than invariant or absolute; (2) in the STR, the spontaneous readings of clocks U which are initially synchronized with the clock U_1 at a space point A will generally *disagree* after transport to another space point B. This intrasystemic discordance among the clocks U makes for an inconsistency between simultaneity verdicts, when the readings of U_1 are paired with the *spontaneous* readings of different clocks U .

Bridgman and also E & B are telling us, in effect, that the second (though not the first!) of these differences between the two worlds can be rendered innocuous. For they point out that the intra-systemic discordance among the spontaneous readings of the clocks U is such as to disappear in the limit of zero velocity or infinitely slow transport in the following sense: in that limit, the reading of U is the same as that of a clock at B which is in standard signal synchrony with U_1 . E & B are careful to note (pp. 129–130) that although a one-way velocity presupposes a clock synchronization, the requirement of approaching infinitely slow transport can be implemented mathematically without invoking some one synchronization to the exclusion of alternative ones.

The clocks in the world of the STR do indeed behave this way *as a matter of physical fact*. Let me pause to show that this is so, since the proof given by E & B (pp. 128–130) needlessly rests on an approximation (p. 128, equation (8)) and also makes avoidable use of an “intervening velocity.”

See Append.
§58 for a
paragraph
missing here

Referring back to our world-line diagram in section 3.1, recall that E_1 is the emission of a light ray at A whose arrival at B coincides with the particular event E' there. Also, as before, E_0 is a particular event at A such that E_0E' is a portion of the world-line of a clock U which moves with respect to our inertial system I along the spatial straight line segment $AB = d$. If the time t_0 of E_0 on the clock U_1 is zero, then the particular events E_0 and E' will determine the later time t_1 of

⁷[6], p. 45. I also mistakenly endorsed this view in several of my subsequent publications. And hence I failed to take issue with Putnam on it when Putnam also espoused it: see [10], pp. 85–86.

E_1 on the clock U_1 at which the light ray departed. That time $t_1 > 0$ on U_1 is given by

$$t_1 = \frac{d}{v_0} - \frac{d}{c},$$

where v_0 is a parameter depending on the particular pair of events E_0, E' , and c is a constant which has the numerical value of the standard velocity of light but does *not*, at this stage, depend on any particular synchronization of U_2 with U_1 . Our concern is with the reading of U at B upon the occurrence of the given event E' there. Clearly, that reading on U will depend on the pair E_0, E' and hence on the parameter v_0 , as well as on the fact that U was *locally* synchronized with U_1 on departing from it. But this U -reading *cannot* depend on whether we have synchronized U_2 with U_1 in standard fashion or not! And nothing that we have said *so far* depends on any kind of synchronism of U_2 with U_1 ! For the *heuristic purpose* of ascertaining the reading on U at B, however, we shall now assume that U_2 is in *standard* synchrony with U_1 .

Given this assumption, the one-way velocity of light is c , the one-way transit time of light for the distance d is d/c , and hence the light emitted at the time $t_1 = d/v_0 - d/c$ on U_1 will reach U_2 when the latter reads $t_2 = d/v_0$. Thus, if we use standard synchronism $\epsilon = 1/2$, then our event-pair parameter v_0 turns out to be the one-way velocity of the clock U , and U_2 will read a time $t_2 = d/v_0$ on the arrival of U . But if we do *not* use standard synchronism, the clock U_2 will read a time different from d/v_0 on the arrival of U , say $d/v_0 + k$ (where $k \neq 0$).

Yet regardless of whether and how we synchronize U_2 with U_1 , the time t' of the event E' on the transported clock U will be

$$t' = \frac{d}{v_0} \sqrt{1 - \frac{v_0^2}{c^2}}.$$

Now consider not just one transported clock but a large number of such clocks U all of which start out in local synchrony with U_1 at various times on that clock and arrive jointly at B upon the occurrence of E' there. If v is an event-pair parameter such that $0 < v \leq v_0$, let these various clocks U depart respectively from A at ever earlier times $t \leq 0$ given by

$$t = \frac{d}{v_0} - \frac{d}{v}.$$

By our earlier reasoning, we can conclude that the time-increment on each clock U for the trip from A to B will be $(d/v)\sqrt{1 - v^2/c^2}$. Hence on jointly arriving at B when the event E' occurs there, the clocks U will read times t' given by

$$t' = \frac{d}{v_0} - \frac{d}{v} + \frac{d}{v} \sqrt{1 - \frac{v^2}{c^2}} = t_2 - \frac{d}{v} \left[1 - \sqrt{1 - \frac{v^2}{c^2}} \right].$$

We wish to determine the limit of these readings as we take values of the event-pair parameter v which correspond to ever earlier times t of departure, i.e. as v approaches zero. If we use $\epsilon = 1/2$ to synchronize U_2 with U_1 , then v is the one-way

velocity of U , and the limit we seek can be characterized as corresponding to infinitely slow transport.

To show that $\lim_{v \rightarrow 0} t' = d/v_0$, where d/v_0 is the standard synchronism reading t_2 on the clock U_2 , we need only show that the *difference* between t' and t_2 vanishes in the limit of $v \rightarrow 0$. Since

$$t' - t_2 = \frac{d \left[\sqrt{1 - \frac{v^2}{c^2}} - 1 \right]}{v},$$

we see that both the numerator $f(v)$ and the denominator $g(v)$ of this fraction approach zero as $v \rightarrow 0$. By l'Hospital's rule,⁸ in this case

$$\lim_{v \rightarrow 0} \frac{f(v)}{g(v)} = \lim_{v \rightarrow 0} \frac{f'(v)}{g'(v)} = \frac{f'(0)}{g'(0)},$$

since $g'(v) = 1 \neq 0$. But

$$f'(v) = \frac{-vd}{c^2 \sqrt{1 - \frac{v^2}{c^2}}},$$

so that $f'(0) = 0$. Hence

$$\lim_{v \rightarrow 0} (t' - t_2) = \frac{0}{1} = 0. \quad \text{Q.E.D.}$$

Thus, the intra-systemic discordance among the spontaneous readings of the clocks U is such as to issue in the unique reading required by standard synchrony, if we go to the limit of infinitely slow transport. Note that all of the time coordinates ingredient in the sequence of differences of which this limit is taken pertain to the one single event E' . For E' is the event at time t_2 on U_2 when all of the clocks U *jointly* arrive at B. This way of taking the limit differs only inessentially, however, from that employed in Bridgman's original method. As will be recalled from the overall Introduction by Salmon and myself, Bridgman's moving clocks C_n do *not* arrive jointly at B. Thus, he takes the limit of a sequence of other differences as follows: the differences between the various readings of the clocks C_n , on the one hand, and the succession of *distinct* times of their arrival on the fixed clock at B, on the other.

The crucial question is what conclusion follows from the physical facts expressed by these limits in regard to the philosophical status of the simultaneity relations corresponding to slow transport synchrony in the STR. Our discussion of the quasi-Newtonian world will now enable us to answer this question, even though slow transport synchrony yields the relative, noninvariant simultaneity relations of the STR, whereas the transport synchrony of the quasi-Newtonian universe issues in absolute relations of simultaneity. We may use the quasi-Newtonian world to

⁸ L'Hospital's rule states the following: If two functions $f(x)$ and $g(x)$ together with their derivatives up to order $(n - 1)$ vanish at $x = a$, and if their derivatives of n th order do not both vanish there or both become infinite, then $\lim_{x \rightarrow a} [f(x)/g(x)] = f^n(a)/g^n(a)$ [see *The International Dictionary of Applied Mathematics* (Princeton, New Jersey: Van Nostrand, 1960), p. 539].

answer our question for the following reason: under *slow* transport, the readings of the clocks in a given inertial system of the STR have the *same intrasystemic physical significance* as the *concordant* readings of the clocks transported in an inertial system of the quasi-Newtonian world. We saw apropos of the latter universe that the mere physical fact of clock concordance under transport is *not* sufficient to assure that the simultaneity relations implicit in clock transport synchrony can be only trivially conventional.

Having rightly noted the *physical fact* that

$$\lim_{v \rightarrow 0} (t' - t_2) = 0,$$

E & *B* infer incorrectly that the time coordinate *t'* which the slowly transported clock assigns to *E'* when it reaches *B* has the following *ordinal* significance with respect to the events at *A*: *as a matter of physical fact, E' occupies the same place in the order of earlier or later relating the events at A as the event E which occurs at the time t = t' on U₁*. But we know from the quasi-Newtonian world that precisely such an inference is fallacious and leads to a false conclusion.

Indeed the analogy between simultaneity and spatial congruence is again instructive. If three rods 1, 2, and 3 coincide with each other at a place *A*, and rods 2 and 3 are then transported to another place *B*, then it is a *physical fact* that the latter two rods will again coincide (in the absence of "differential" forces) independently of their paths of transport. This physical law assures the *consistency* of the statement that *all* (differentially undeformed) rods remain self-congruent under transport. But, as we learned from Riemann, it does *not* assure that this assertion of self-congruence can be only *trivially* conventional. The nontrivial conventionality of the self-congruence of the class of rods is not to be doubted even though this assertion presupposes a law of *concordance* among the congruence findings of different rods, which holds as a matter of physical fact. To be sure, the self-congruence of the class of rods would become a factual matter relatively to another standard of spatial congruence such as the *round-trip* time of light. But then the conventionality creeps in via the self-congruence of *that* standard under transport.⁹

By the same token, it is unavailing for *E* & *B* to point out that the behavior of clocks under slow transport is a uniquely specified matter of fact, just because it is a fact that $\lim_{v \rightarrow 0} (t' - t_2) = 0$. For this fact does not establish that the simultaneity relations corresponding to slow transport synchrony are matters of ordinal temporal fact and only trivially conventional. Thus these simultaneity relations do not become only trivially conventional just because it is a fact that clocks which are in slow transport synchrony are also in standard signal synchrony. The latter concordance does show that both slow transport and Einstein's light signal rule will *alike* yield a time coordinate for *E'* which is equal to the coordinate *t* yielded by the clock *U₁* for the event *E*. Hereafter, let us denote the time coordinate of *E'* which is furnished by the *slowly* transported clock by *t'_s*. In other words, $\lim_{v \rightarrow 0} t' = t'_s$. Then it is a fact that *t'_s = t₂*, where *t₂ = t*. Furthermore, it is a fact that *E* and *E'*

⁹ For a very detailed discussion of this point and of its ramifications, see [10], §§2.9, 2.10, 2.11, and §6.

are both temporally between E_1 and E_3 . But, nevertheless, without an appeal to a time metric, the combination of these facts is *not* sufficient to remedy the following aforementioned hiatus in the temporal order: there is no physical basis here for asserting that E' is also temporally between pairs of events x and y such that x belongs to the open interval E_1E , and y to the open interval EE_3 .

Since the ordinal hiatus needs to be filled here by a metrical claim in order to assert simultaneity, the legitimacy of alternative durational metrics makes for the legitimacy of corresponding alternative relations of simultaneity. In particular, despite the equality of the time coordinates of E and E' , the resulting equality of the differences between the time coordinates of the pairs E_1E and E_1E' does *not* compel us to assign equal durational measures to these two event pairs, and hence does not compel us to assert the simultaneity of E' and E . By the same token the equality of the differences between the time coordinates of the pairs EE_3 and $E'E_3$ does not compel us to make this assertion of simultaneity. In particular, suppose that we let the durational measure of event intervals such as E_1E on the world-line of U_1 be given by the difference $t - t_1$ between their time coordinates in the manner of the standard time metric. Then we are free to let the durational measure of the interval E_1E' on the world-line of the outgoing light ray be given by the quantity $2\epsilon(t'_s - t_1)$, where ϵ may be *any* value between 0 and 1 and hence may differ from $\frac{1}{2}$. And we are likewise free to let the durational measure of the interval $E'E_3$ be given by the quantity $2(1 - \epsilon)(t_3 - t'_s)$. Note that when following *this* procedure to introduce a nonstandard time-metric by a choice of $\epsilon \neq \frac{1}{2}$, we do *not* tamper with the setting t'_s of the slowly transported clock. Therefore our particular procedure here does not call for imparting a setting to this clock which differs from the standard signal synchrony reading t_2 of the clock U_2 . Nevertheless, our procedure here issues in the same simultaneity verdicts as *non-standard* signal synchrony, because we are doing the following: (1) we are accepting the time coordinate $t'_s = t$ of the slowly transported clock but employing the specified *non-standard* time metrics on the world-lines of the outgoing and returning light rays, which connect E' to events on U_1 , and (2) we are employing the standard time metric on the world-line of U_1 . In short, here we avail ourselves of our freedom to deny that the time coordinates furnished by the slowly transported clock automatically qualify it to be *synchronized* with U_1 .

Using $\epsilon \neq \frac{1}{2}$ in the context of the *latter* procedure, let us determine what event E_x on U_1 other than E turns out to be simultaneous with E' . This event E_x is now specified by the fact that E_x and E' must be durationally equidistant from E_1 , provided that the durational measure of E_1E_x is given by the standard time metric, while the durational measure of E_1E' is given by the nonstandard time metric $2\epsilon(t'_s - t_1)$. Let t_x be the time coordinate of E_x on U_1 . Then t_x is given by the equidistance condition

$$t_x - t_1 = 2\epsilon(t'_s - t_1).$$

Recalling that

$$t_1 = \frac{d}{v_0} - \frac{d}{c}$$

and

$$t'_1 = \frac{d}{v_0},$$

we obtain for the time coordinate of E_x on U_1

$$t_x = \frac{d}{v_0} \left[1 - \frac{v_0}{c} (1 - 2\epsilon) \right].$$

The same event E_x turns out to be simultaneous with E' , if we use *non-standard* signal synchrony in the context of the standard time metric and set the clock U_2 in accord with $\epsilon \neq 1/2$. For *this* procedure issues in a time coordinate t'_2 for E' different from t_2 and given by

$$t'_2 = t_1 + \epsilon \left(\frac{2d}{c} \right).$$

Substituting for t_1 , we obtain

$$t'_2 = t_x.$$

Incidentally, in order to obtain the simultaneity of E' and E_x via their durational equidistance from events such as E_0 , (i.e. from events before E_1), the durational measure of E_0E' must be given by a *more complicated* metric than in the case of E_1E' .

Let us now apply the nonstandard time metrics $2\epsilon(t'_2 - t_1)$ and $2(1 - \epsilon)(t_3 - t'_2)$ to the time coordinates furnished by U_1 and by the slowly transported clock to determine the one-way velocities of light along AB and BA. We then obtain

$$v_{AB} = \frac{d}{2\epsilon(t'_2 - t_1)} = \frac{d}{2\epsilon \cdot d/c} = \frac{c}{2\epsilon}$$

and

$$v_{BA} = \frac{d}{2(1 - \epsilon)(t_3 - t'_2)} = \frac{d}{2(1 - \epsilon)(d/c)} = \frac{c}{2(1 - \epsilon)}.$$

And these to and fro velocities are identically those obtained from nonstandard signal synchrony $\epsilon \neq 1/2$. Thus, the facts of slow transport behavior of clocks, coupled with the two-way light principle, cannot compel us to assert the equality of the to and fro velocities of light, any more than other facts in the STR can compel us to employ standard signal synchrony and thereby to assert that equality. For we saw that the time coordinate furnished by the slowly transported clock cannot remove the aforementioned ordinal hiatus, cannot dictate a time metric, and cannot confer factual truth on the particular time metric employed to infer simultaneity. In other words, the facts of slow transport behavior of clocks do *not* confer factual truth on the assertion of the simultaneity relations corresponding to slow transport synchrony. And these simultaneity relations are just as non-trivially conventional—hereafter “Riemann conventional”—as the simultaneity relations implicit in standard signal synchrony. But, of course, *if* we do adopt slow transport

synchrony to *stipulate* simultaneity, then the *equality* of the to and fro velocities of light becomes an empirical matter of fact. By contrast, if we adopt standard signal synchrony to stipulate simultaneity, then the equality of these *one-way* velocities is a mere *definitional* consequence of the synchronization rule. The adoption of slow transport synchrony confers *factual truth* on the assertion of this velocity equality for the simple reason that the agreement between slow transport and standard signal synchrony is factual in the STR. But the fact of this agreement cannot nullify the Riemann-conventionality of the simultaneity relations implicit alike in each of the concordant synchronies, and in the equal light velocities.

This conclusion invalidates *E* & *B*'s characterization of the *philosophical significance* of Römer's determination of the one-way velocity of light reaching the earth from the moons of Jupiter. Reichenbach had pointed out that the synchronism implicit in Römer's determination is furnished by the earth as a transported clock ([15], pp. 628–631). And, in concert with Reichenbach, I had stated that here as elsewhere in the STR, "no statement concerning a one-way transit time or one-way velocity derives its meaning from mere facts but also requires a prior *stipulation* of the criterion of clock synchronization" ([9], p. 355, and p. 355, n. 6). *E* & *B* first comment (p. 131) that "Römer's method, in particular, is a legitimate procedure for determining the one-way velocity of light." And in a footnote (p. 131, n. 16), they say in part:

Grünbaum, in citing Reichenbach, interprets the problem as being one of tacitly assuming "a *prior stipulation* of the criterion of clock synchronization". . . . However, we view the terrestrial clock as being *known* empirically to be in slow transport synchrony. According to this view the synchrony in question is not given by fiat, as it is with the previous authors; but rather it is known to be true directly and independently of signalling procedures. This being known, the one-way velocity of light can be measured.

But we just saw that though it is empirical, Römer's determination of the one-way velocity of light is *not* based on a slow transport synchrony which "is known to be true."

Suppose that, *contrary to fact*, Römer had found the velocity of the light from Jupiter's moons to depend upon the direction of Jupiter in the solar system by using the slow transport synchrony of the terrestrial clock. In that hypothetical eventuality, there would be disagreement between slow transport and standard signal synchrony, *contrary* to the STR. But suppose further that in conformity to the STR, we have every empirical reason to assume that light is the fastest possible causal chain and that its round-trip velocity is positive. Call this hypothetical universe "World #1." Then in the face of the above *hypothetical disagreement* between the two synchronies in World #1, the facts of slow clock transport would *not* compel us inductively to interpret Römer's *hypothetical* findings as establishing the inequality of the to and fro velocities of light. Precisely because it is *not* the case that slow transport synchrony "is known to be true," we could interpret the putative findings as allowing the equality of the to and fro velocities of light.

Specifically, suppose that, contrary to fact, Römer's slowly transported terrestrial clock had found that the light ray from A arrived at B not at the time $t'_s = d/v_0$, but at a time $t'_x = d/v_0 - (d/c)(1 - 2\epsilon)$, where the value $\epsilon \neq 1/2$ is between 0 and 1. Let $\delta \equiv (d/c)(2\epsilon - 1)$, so that we can write $t'_x = d/v_0 + \delta$. And be mindful of the

nonexistence of facts that would compel us to regard the simultaneity implicit in slow transport synchrony as "true." Then, in World #1, we are free to assure the equality of the to and fro velocities of light by doing the following: choosing time-metrics on the world-lines of the to and fro light rays that would assign equal durational measures to the event pairs whose coordinate differences are now presumed to be $t'_x - t_1$ and $t_3 - t'_x$. This can be done by choosing any two nonvanishing multipliers M and N such that the outgoing transit time of light $M(t'_x - t_1)$ is equal to the return transit time $N(t_3 - t'_x)$. And it is clear that the requisite temporal equidistance condition

$$M\left(\frac{d}{c} + \delta\right) = N\left(\frac{d}{c} - \delta\right)$$

can easily be fulfilled by $M = 1/(2\epsilon)$ and $N = 1/[2(1 - \epsilon)]$, which render E' simultaneous with E .

There is a *different* modification of the universe of the STR which makes further apparent that the simultaneity implicit in slow transport synchrony as such is not "true" in the STR in a sense incompatible with being Riemann-conventional. For suppose now that contrary to the STR, there are arbitrarily fast causal chains whose properties are those of the members of the class K of causal chains *other than clocks* in Newton's world (cf. section 3.1). In that case, these causal chains define absolute and nonconventional relations of simultaneity, which preclude relative, frame-dependent relations of simultaneity. But assume further in conformity to the STR that (1) no transported material clock can keep pace with a light ray, (2) clocks in slow transport synchrony in the various inertial systems yield the relativity of simultaneity required by the Lorentz transformations, and (3) there is agreement between slow transport and standard light signal synchrony. We shall call this second hypothetical universe "World #2." It will be recalled that the numbers furnished by (moving) clocks derive their original significance as time coordinates on their own world-lines from the following fact: the career of the center of mass of a clock is a causal chain, and the relations of *numerical* betweenness among the coordinates are isomorphic with the relations of *causal* betweenness among the events to which the coordinates are assigned by the clock. But clock careers constitute only *one* of the species of causal chain. Yet in our hypothetical modified World #2, no less than in the unmodified universe of the STR, the facts of slow clock transport and of light propagation do leave the aforementioned ordinal hiatus in the time-system, *unless* they are coupled with the facts pertaining to super-light causal chains. For precisely this reason, the putative existence of the latter causal chains would foreclose the option of asserting the *relative*, noninvariant simultaneity relations implicit in slow transport synchrony *and* in standard light signal synchrony.

Thus, in our putative World #2, the arbitrarily fast causal chains would constitute a physical basis for relations of *absolute* simultaneity which are *not* Riemann-conventional. And hence these chains would *rule out* the adoption of the conventions that give rise to the relative, discordant simultaneity verdicts implicit in slow transport synchrony and in standard light signal synchrony. Under the hypothetical

circumstances of World #2, slow transport synchrony would issue in false assertions of simultaneity. The status of simultaneity in World #2 serves to refute *E* & *B*'s claim (p. 131, n. 16) that in the unmodified universe of the STR, slow transport synchrony "is known to be true." And it does so partly because slow transport synchrony does not rest in either universe on the simultaneity relations implicit in any synchrony based on the class of causal chains *other than* moving clocks.

In the context of the particular absolute simultaneity of World #2, it would be a *matter of fact* that the to and fro velocities of light are unequal in some inertial systems. By contrast, we saw that it would *not* be a matter of fact that these one-way velocities are unequal in the inertial systems of World #1, though the adoption of slow transport synchrony in the latter world would render them unequal.

Having mistakenly claimed that slow transport synchrony "is known to be true" (p. 131, n. 16), *E* & *B* draw two conclusions as follows (p. 131 and p. 131, n. 17): (1) They assert incorrectly that the to and fro velocities of light in World #1 are unequal as a matter of fact, unalloyed by Riemann-conventionality, (2) they erroneously reject my contention that World #2 presents us with "physical facts incompatible with the conventionality of simultaneity" *only* because it contains super-light signals *and* that these signals would make for "the illegitimacy" of asserting the equality of the to and fro velocities of light in every inertial system.¹⁰ And they maintain that as between Worlds #1 and #2, only World #1 presents us with such facts. Having drawn this false conclusion, they comment on my characterization of World #2 by saying (p. 131, n. 17):

... we do not understand how such a statement can be compatible with his conventionality thesis, especially this statement: "... no distinct hypothesis concerning physical facts is made by the choice of $\epsilon = \frac{1}{2}$ as against one of the other permissible values."

Their puzzlement is of their own creation, since my *latter* statement pertains to the world of the unmodified STR, whereas my former contention refers to the quite different World #2. The incompatibility here lies *not* in my cited account of the conditions under which simultaneity is conventional; instead the incompatibility obtains between the two different worlds which I was discussing. Of course, the phrase "the choice of $\epsilon = \frac{1}{2}$ " in my statement refers to the stipulation of simultaneity implicit in *both* slow transport and standard signal synchrony, *not* to the factual *agreement* between these two synchronies.

Furthermore, *E* & *B* (p. 125) misconstrue me as having maintained the following: the existence of simultaneity relations in the STR is predicated on a prior specification of a one-way velocity, and this dependence issues in a vicious circle which can be broken only by introducing simultaneity as a matter of convention. I had indeed maintained that a determinate numerical one-way velocity does presuppose distant simultaneity for the reasons stated at the end of 3.1 above ([5], pp. 172–173). But the context in which I discussed the *inverse* presupposition, in the publications cited by *E* & *B* in this connection, was the following: a single observer *O* at a given place *C* is presented with the local intersection at *C* of causal chains that have

¹⁰ The relevant statement of mine is given in [9], p. 396, item (2) and p. 397, item (2). *E* & *B* misparaphrase my statement by representing me as saying that the infinitely fast causal chains in World #2 "yield $\epsilon = \frac{1}{2}$." And there are distorting omissions of primes in the quotation.

emanated from two other separate space points A and B. And the problem is one of the logical presuppositions of information that would then permit the single observer *O* at C to *infer* the simultaneous origination of the causal chains at A and B respectively.

Thus, speaking of Whitehead's "thought that distant simultaneity can be readily grounded on local simultaneity," I pointed out the following: the procedure of inferring distant simultaneity from the one-way velocities of the causal chains (and from the respective distances of A and B) "is unavailing for the purpose of first characterizing the conditions under which the separated events E_1 and E_2 can be held to be simultaneous. For the *one-way* velocities invoked by this procedure presuppose *one-way transit times* which are furnished by *synchronized* clocks at the locations of E_1 and E_2 , respectively. And the conditions for the synchronism of two spatially separated clocks, in turn, presuppose a criterion for the distant simultaneity of the events at these clocks" ([9], p. 344; see also [11], p. 224). In the universe of the STR, the conclusion of my cited argument here holds as much for slow transport synchrony as for standard or nonstandard light signal synchrony. Indeed, we saw that it holds no less in the Newtonian and quasi-Newtonian worlds. For a given physical theory regards clocks as synchronized if they read the same time numbers for events which qualify as simultaneous within the given physical theory. Moreover, the existence of slow transport synchrony in the STR also accords entirely with the *premises* of my argument here. *E* & *B* point out illuminatingly that the feasibility of slow transport synchrony does refute the claim that standard distant simultaneity can be specified in the STR *only* by a one-way velocity or *only* by the *equality* of the to and fro velocities of light. But I had not asserted *this* dependence apropos of standard signal synchrony. Nor had I rested the conventionality of simultaneity on it.

Reichenbach did affirm that in the absence of an initial stipulation of simultaneity, simultaneity presupposes a one-way velocity and is thereby beset by circularity ([14], pp. 124–128). But it is unclear that he intended this to be an *unqualified* claim, as *E* & *B* charge (pp. 125–126, and 131). If he did so in the place cited by *E* & *B* ([14], 126 ff.), then he was being inconsistent. For only some pages later ([14], p. 129 and pp. 146–147), he chides "certain relativistic presentations" ([14], p. 146) for having erroneously maintained that the notion of a *non*-arbitrary ([14], p. 129) and invariant ([14], pp. 146–147) simultaneity relation "is a logically impossible conception" ([14], p. 146).

As for Einstein's own view on the dependence of simultaneity on velocity, Perrett and Jeffery unfortunately mistranslated the pertinent sentence in his original 1905 German paper into English ([3], p. 40). Hence their mistranslation can be mistaken as documentary proof that Einstein himself erroneously asserted that simultaneity can be specified in the STR *only* by stipulating the equality of the to and fro velocities of light. But in Einstein's original German text, he had offered the stipulation of the equality of these one-way velocities as a (contextually) *sufficient* condition for simultaneity, and *not* as a necessary condition ([16], p. 398).

It remains to deal with two further arguments given by *E* & *B* against our thesis

that the simultaneity ingredient in slow transport synchrony is Riemann-conventional. They write (p. 126):

Reichenbach claims, moreover, that even if . . . it were possible to establish a synchrony by a transport procedure, it would still be a matter of definition, and not of empirical fact, that clocks that are locally synchronous remain synchronous after separation. For the case would then be entirely analogous to that of distant congruence. If we were prepared to allow that there be "universal forces" that act equally on all clocks, and cannot be insulated against, then . . .

(pp. 134–135):

. . . if a nonstandard slow transport synchronization were accepted which was consistent with the n -way light principle and the topological condition for light, we should have to introduce two sets of universal forces. One set would be needed to account for the anisotropy of the empirically determined one-way velocities of light, and a second set would be needed to account for the acceleration or retardation of infinitely slowly transported clocks. They would be universal forces in the sense that they affect all electromagnetic signals or all clocks equally, and in the sense that they cannot be insulated against. Moreover, the two sets of universal forces would have to have identical distribution laws, i.e. they would have to be laws identical in form to the distribution law for light velocities. Hence, unless Reichenbach's conventionality thesis is the trivial one mentioned above (in which case it is not worth discussing), Reichenbach's conventionality thesis is false. The existence of two logically independent ways of defining distant simultaneity each of which is unique if the principle of setting universal forces to zero is accepted and which, as a matter of empirical fact, appear to define the same relationship of distant simultaneity, provides a good physical reason for adopting one of these definitions. If physical reasons such as these are ruled out of court then there can be no good physical reason for adopting any definition of any quantitative equality between any two things at a distance. Hence, if the predictions of the Special Theory of Relativity regarding transported clocks are correct, there are good physical reasons for adopting a definition of synchrony which either sets $\epsilon_0 = 1/2$ by definition, or from which $\epsilon_0 = 1/2$ can be shown empirically to be true.

This passage contains two sets of contentions as follows:

(1) Reichenbach's conventionality thesis is either false or only trivially true. In its nontrivial version, it is false because it requires the introduction of *two* independent sets of universal forces which obey identical distribution laws. And—as a matter of empirical fact—these forces do not exist.

(2) If the principle of setting universal forces to zero is accepted, then there are "good physical reasons" for $\epsilon = 1/2$. And hence simultaneity is not Riemann-conventional but only trivially so.

These contentions prompt two corresponding sets of comments.

1. As we explained earlier on the basis of Riemann's ideas, we are free to choose any time metrics $2\epsilon(t'_s - t_1)$ and $2(1 - \epsilon)(t_3 - t'_s)$ in which $0 < \epsilon < 1$ as the durational measures of the light ray event intervals E_1E' and $E'E_3$ respectively. And if we avail ourselves of this freedom to adopt a *non*-standard time metric ($\epsilon \neq 1/2$), then E' becomes simultaneous with the event E_x whose U_1 coordinate is t_x . On the other hand, again given that the time coordinate of E' is t'_s on the slowly transported clock, we may adopt the standard time metric on these optical world-lines ($\epsilon = 1/2$). And if we do, then E' becomes simultaneous with E instead of with E_x . On the former choice, the to and fro velocities of light are *automatically* rendered unequal, whereas the latter choice automatically renders them equal.

At the risk of being misleading, one can invoke "universal forces" *metaphorically*

(!) to characterize choices among alternative metrics.¹¹ In particular, one can take this risk here by describing the choice of our *non-standard* time metric in terms of the language of the standard time metric. This somewhat confusing description takes the following form. One says first that E_1' and E_3' (rather than E) are durationally equidistant from E_1 , and that the durational measure of E_1E_3' *should* really be given by the *mere difference* of the time coordinates just like the measure of E_1E_3 . Then one says that the slowly transported clock should therefore have exhibited this equidistance by reading a time coordinate $t_3' = t_1$ instead of having read $t_3' = t$. And one can go on to declare *metaphorically* that the slowly transported clock would indeed have read t_3' *for the very same optical event E_3'* , if it had not been under the influence of "universal forces" which "caused" it to read $t_3' = t$ instead. This declaration has *no physical content* and is merely a pictorial device to characterize the *non-standard* time metric, which is applied to the *ACTUAL* time coordinates furnished by the clocks. And these actual coordinates differ by d/c as a matter of fact. Thus, it is a mere metaphor to talk here in this way about the action of universal forces on the slowly transported clock in the interest of nonstandard metrics.

But we showed that the use of the above nonstandard time metrics *automatically* assures the inequality of the to and fro velocities of light. For this inequality is merely a matter of the nonstandard durational measures of the given pairs of optical events E_1E_3' and $E_1'E_3$. Hence even if one misconstrues Reichenbach's metaphor of "universal forces" here with E & B as literal, it is altogether gratuitous on their part to assert the following: *two independent sets* of universal forces, which happen to satisfy identical distribution patterns, are needed here to describe the *identical optical phenomena* by means of non-standard slow transport synchrony. These optical phenomena consist in the fact that the slowly transported clock reads $t_3' = t = d/v_0$ upon the arrival at B of a light ray which departed from A at the time $t_1 = d/v_0 - d/c$ and returns to A at $t_3 = d/v_0 + d/c$.

This fact can be described by saying that the slow transport and standard signal synchronies agree, so that the to and fro light velocities are equal. It can be so described on the strength of having used the *standard* time metric on the optical world-lines. This metric yields the one simultaneity convention implicit alike in each of these two synchronies. Alternatively, the optical phenomena can be described by applying nonstandard time metrics to the *identical* time coordinates furnished by the two clocks for the events on the optical world-lines. In the latter case, there is automatic inequality of the to and fro light velocities. *It is merely a matter of the use of alternative time metrics on these world-lines.* Whatever durational metrics we use for the intervals E_1E_3' and $E_1'E_3$, whose coordinate differences are each in fact d/c , we *thereby* determine the "absence" or "presence" of "universal forces."

Therefore, within the framework of the STR, it is no more incumbent upon *one* of these metrical descriptions to "account" for the *postulated* fact described by it than upon the other. Within the framework of the STR, not even *one* set of "universal forces" *in the literal sense* is needed by the *non-standard* description. *E* & *B*

¹¹ I discussed this point in detail in [9], Ch. 3, Section A, where I warned against confusing the metaphorical and literal uses of "universal forces."

address the demand to “account” for the fact postulated by the STR *only* to those who claim the legitimacy of the nonstandard time metric. Their indictment of the allegedly required particular account as false (*ad hoc*) is therefore a straw man. And so is the inductive implausibility of a pre-established harmony between the two sets of universal forces which are allegedly needed to provide such an account.

2. Earlier in the present section 3.3, we discussed the spatial congruence relations specified by the class of solid rods which are conventionally self-congruent under transport. We saw that this customary convention of self-congruence presupposes the fact of concordance in the coincidence behavior of the rods, i.e. that this concordance is a *necessary* condition for the *consistency* of the conventional self-congruence of *all* rods. But the fact of concordance is only a necessary and not also a sufficient condition for the customary convention. For the concordance also allows and indeed is needed to assure the consistency of an *alternative* congruence convention according to which the length of *any* transported rod is the *same* non-constant function of its position and/or orientation! In short, the fact of concordance *entails* neither the standard congruence convention nor any one of the non-standard ones. And this concordance could be held to refute the claim that the customary congruence is Riemann-conventional, i.e. has physically allowable alternatives, *only* if the concordance did entail that one congruence to the exclusion of the others.

Hence the fact of concordance is “a physical reason” for the *customary* congruence convention in the sense of being a necessary physical condition for it. But it is *not* “a physical reason” for it in the *different* sense of impugning the Riemann-conventionality of that particular congruence specification by disallowing an alternative one. And it would therefore be most wrongheaded to say that those who assert the Riemann-conventionality of congruence are ruling “out of court” physical reasons such as this concordance! What is being ruled out of court is only the fallacious reasoning which adduces the concordance in an attempt to establish that the self-congruence of the class of rods is a matter of spatial fact.

By the same token, the fact of clock behavior under slow transport is a *necessary* but not a sufficient condition for the simultaneity relations implicit in slow transport synchrony. For the coordinate t'_s of E' , which is furnished by the slowly transported clock, also allows alternative simultaneity relations. Hence the physical fact of slow transport behavior is not ruled out of court by the thesis that simultaneity is Riemann-conventional. What this thesis does rule out, however, is E & B 's fallacious inference that the physical fact in question establishes the *physical truth* of one of the two concordant standard synchronies to the exclusion of nonstandard ones. Thus E & B are mistaken when they claim the following (p. 127): “the sense in which distant simultaneity could be said to be conventional would be both interesting and important” *only* if light signal connectibility were the *sole* physical relationship “which could define a relationship of synchrony.”

Let us now review the philosophical status of simultaneity in the STR by considering (a) the situation in the absence of reliance on slow transport synchrony, and (b) the altered situation made possible by the availability of both slow transport and light signal synchrony.

(a) *Simultaneity by Light Signal Synchrony Alone*

The events belonging to the world-line of a single clock are ordered by a relation of causal betweenness. And this order of betweenness is *reflected* by the time coordinates which are *spontaneously* furnished by the clock, provided that the numerals painted on the dial of the clock have the appropriate *spatial* order. If the numerals had been painted spatially at random, then the time coordinates would no longer reflect the relations of causal betweenness, but without detriment to the existence of these relations. Thus the time coordinates furnished by a *suitably painted* clock dial derive their *ordinal significance* from the objective relations of causal betweenness. But it was *our* decision to paint the dial in a certain way, which permitted the resulting time coordinates to reflect this objective betweenness. This is the case for our events E_1 , E and E_3 on U_1 , for example.

But the three events $E_1E'E_3$ cannot all belong to the career of a single clock. Hence no single clock can furnish time coordinates which would reflect the objective relation of causal betweenness that exists among them on the world-line of the light ray. If this relation is to be reflected at all by time coordinates which are furnished by clocks, it will require *two* different clocks which are suitably positioned. Just as it was our decision to paint the clock dial of U_1 in the requisite way, so also it is up to us to *impart a setting* to a second clock U_2 which satisfies the following requirement: the resulting time coordinate t_2 of E' on U_2 , coupled with the time coordinates t_1 and t_3 on U_1 are to reflect the objective causal *betweenness* which relates E' to E_1 and E_3 on the world-line of the light ray.

This requirement yields the topological condition for light

$$0 < \epsilon < 1.$$

Note incidentally that this condition for light does *not* follow from the coordinative definition of temporal separation or nonsimultaneity (our "N-S" in 3.1 above). The latter coordinative definition merely calls for the assignment of different time numbers by clocks to *any* two events which sustain the *symmetric* relation of causal connectibility. E_1 , E' and E_3 are pairwise causally connectible. Hence the pairwise application of coordinative definition N-S requires merely that t_2 *differs* from both t_1 and t_3 . But this obviously allows t_2 to be less than t_1 or greater than t_3 , i.e. it allows $\epsilon < 0$ or $\epsilon > 1$. In short, the coordinative definition N-S does not utilize the causal *betweenness* relation among the three events to confine t_2 to values between t_1 and t_3 as required by the topological condition for light.

The topological condition for light stipulates that U_2 is to be set so as to assign to E' any coordinate t_2 which is between t_1 and t_3 . And E & B agree that here the residual ordinal hiatus *as between E' and the events in the open interval E_1E_3* leaves scope for a conventional stipulation of metrical simultaneity.

(b) *Simultaneity by Slow Transport Synchrony*

E & B are perfectly correct in emphasizing that the slowly transported clock *presents us* with a *unique* time coordinate t'_s for the optical event E' , whereas a light ray itself delivers no such coordinate, unless we *stipulate* it. They tell us further, however, that the existence of the nonarbitrary coordinate t'_s constitutes a "physical

reason" for regarding E' as uniquely simultaneous with E as a matter of ordinal temporal fact. And if the latter claim were granted, then to impart the setting $t_2 = t'_1$ to U_2 upon the occurrence of E' would be merely a matter of letting U_2 exhibit this factually "true" simultaneity. Thus, E & B contend that slow transport behavior does indeed remove the ordinal hiatus left by the topological condition for light. But our analysis has endeavored to show in detail that precisely this central conclusion of theirs is unsound. Hence they have failed to refute the thesis that metrical simultaneity in the STR is Riemann-conventional.

In noninertial frames such as a rotating disk, *standard light signal synchrony generally fails*. And, in general, in such frames even the topological condition for light is *not* a requirement governing the time coordinatization.¹² This raises the question whether *slow transport synchrony* is feasible in the noninertial frames of the GTR. The paper by A. Janis below will examine this question. Janis will show that there is a large class of non-inertial frames, which includes the rotating disk in flat space-time, in which slow transport synchrony fails. Hence there would be scope for conventional choices of simultaneity in all of these cases, even if E & B had given a correct philosophical characterization of simultaneity by slow transport synchrony.

REFERENCES

- [1] Bergmann, P., *Introduction to the Theory of Relativity*, Prentice-Hall, Inc., New York, 1946.
- [2] Bridgman, P. W., *A Sophisticate's Primer of Relativity*, Wesleyan University Press, Middletown, 1962.
- [3] Einstein, A., "On the Electrodynamics of Moving Bodies," *The Principle of Relativity, A Collection of Original Memoirs* (ed. by A. Sommerfeld, tr. by W. Perrett and G. B. Jeffery), Dover Publications, New York, 1952.
- [4] Ellis, B. and Bowman, P., "Conventionality in Distant Simultaneity," *Philosophy of Science*, vol. 34, 1967, pp. 116-136.
- [5] Grünbaum, A., *Epilogue* in P. W. Bridgman's *A Sophisticate's Primer of Relativity*, Wesleyan University Press, Middletown, 1962.
- [6] Grünbaum, A., "The Genesis of the Special Theory of Relativity," *Current Issues in the Philosophy of Science* (ed. by H. Feigl and G. Maxwell), Holt, Rinehart and Winston, New York, 1961.
- [7] Grünbaum, A., *Geometry and Chronometry in Philosophical Perspective*, University of Minnesota Press, Minneapolis, 1968.
- [8] Grünbaum, A., *Modern Science and Zeno's Paradoxes*, Wesleyan University Press, Middletown, 1967, and Allen & Unwin, London, revised edition, 1968.
- *[9] Grünbaum, A., *Philosophical Problems of Space and Time*, Alfred A. Knopf, New York, 1963.
- [10] Grünbaum, A., "Reply to Hilary Putnam's 'An Examination of Grünbaum's Philosophy of Geometry,'" *Boston Studies in the Philosophy of Science*, vol. V (ed. by R. S. Cohen and M. Wartofsky), D. Reidel Publishing Company, Dordrecht, 1969.
- [11] Grünbaum, A., "Whitehead's Philosophy of Science," *The Philosophical Review*, vol. LXXI, 1962, pp. 218-229.
- [12] Mehlberg, H., "Relativity and the Atom," *Mind, Matter, and Method: Essays in Philosophy and Science in Honor of Herbert Feigl* (ed. by P. K. Feyerabend and G. Maxwell), University of Minnesota Press, Minneapolis, 1966.
- [13] Reichenbach, H., *Axiomatik der relativistischen Raum-Zeit-Lehre*, F. Vieweg & Sons, Braunschweig, 1924. (An English translation by Maria Reichenbach with a Foreword by Wesley Salmon will shortly be published by the University of California Press.)

¹² For an illustration of these points and an account of some of the reasons for them, see [7], Ch. III, §§2.3, 2.4, and 8.2.

- [14] Reichenbach, H., *The Philosophy of Space and Time*, Dover Publications, New York, 1958.
- [15] Reichenbach, H., "Planetenuhr und Einsteinsche Gleichzeitigkeit," *Zeitschrift für Physik*, vol. 33, 1925, pp. 623-631.
- [16] Scribner, C., "Mistranslation of a Passage in Einstein's Original Paper on Relativity," *American Journal of Physics*, vol. 31, 1963, p. 318.
- [17] Tolman, R. C., *Relativity, Thermodynamics and Cosmology*, Oxford University Press, Oxford, 1934.

* Parts I-III of the present volume

THE BEARING OF PHILOSOPHY ON THE HISTORY OF THE SPECIAL THEORY OF RELATIVITY*

*Philosophical Mastery of the Special Theory of Relativity is Required for
Unraveling Its History*

In what precise ways is philosophy instrumental in illuminating the *genesis* of the conceptual innovations wrought by a particular physical theory? In the first edition (1963) of this book and in some still earlier papers, I have used the unraveling of the history of the special theory of relativity to argue concretely that philosophy does have far-reaching relevance to the attainment of the following cardinal objectives of the historian of science: (i) the very posing of well-conceived, searching historical questions and (ii) the avoidance of serious historical blunders of certain kinds, and their discernment as such when they have been committed by those lacking the requisite philosophical mastery ([1], chap. 12; [2]). Specifically, I maintained in the context of the special theory of relativity that there is a symbiosis of the philosophy and the history of science as follows: no historically correct, let alone illuminating account of the development of that theory can be furnished without a prior rigorous comprehension of the philosophical conceptions underlying it and distinguishing it from its ancestors. At the same time, I recognized that the history of the theory, in its turn, may indeed contribute to the philosophical analysis of the theory by disclosing the vicissitudes in Einstein's own philosophical outlook.

I now return to the theme of these earlier publications in order to develop it anew. And I do so for the following reasons: (i) I can now supply explicit and specific support for my thesis from Einstein himself, in the form of source materials. The materials in question were either published since or had been previously unknown both to me and presumably to nearly all interested people. (ii) I have had second thoughts on my earlier logical *and* historical assessment of the charge that the *aether-theoretic* Lorentz-Fitzgerald contraction hypothesis and the *aether-theoretic* Lorentz-Larmor time-dilation hypothesis were severally and collectively *ad hoc*. My revised assessment of the charges against these auxiliary hypotheses is prompted by recognition of the need for making previously neglected distinctions between quite different senses in which

a collateral hypothesis can have the logical status of being *ad hoc*. If this revised analysis is sound, it will have quite *general* relevance to the philosophy and history of science.

The source materials which have induced me to provide a fresh treatment of some facets of my earlier theme consist of two reports of interviews with Einstein in which he was asked to recall in as much detail as possible the thought processes which led him to propound the special theory of relativity. The first of these came to light in 1963 and 1973 with the publication of R. S. Shankland's *Conversations with Albert Einstein*, which constitutes a record of what Einstein recalled during 1950 to 1954 concerning the genesis of the theory [3]. The second report, which I ran across late in 1963, is an account by the Gestalt psychologist Max Wertheimer of conversations he had with Einstein starting in 1916 concerning *The Thinking that led to the Theory of Relativity* [4, 5].

I welcomed the opportunity of first taking this further philosophical glance at the history of the special theory of relativity in an address in Cleveland. For it was in that city that Michelson, Morley, and Miller carried out interferometric investigations which figured prominently in the history of the debate on the theory. Cleveland was also the site at which in 1888 Michelson gave his vice-presidential address on optical research to the physics section of the AAAS. It is an irony of history that in that address of 1888 Michelson saw fit *not* to mention the now celebrated null result which he and Morley had obtained in 1887 [6], a result which the Cleveland scientific community of that time regarded as evidence for the *failure* of that experiment and which led to Morley's being 'an object of pity' ([7], p. 55; p. 49, note 9). Perhaps Michelson's lack of pride in the null outcome of the experiment was accompanied by a premonition that it might inspire theoretical developments in physics that would prove most unpalatable to him. For on several occasions Michelson honestly and good-naturedly expressed his distaste to Einstein personally for the theories that had been ushered in by his optical work ([7], p. 57). In fact, Michelson's dislike of the relativity theory was such that he referred to it in Einstein's presence as a 'monster' ([7], p. 56). Similarly uncomplimentary characterizations of the theory are given in a 1951 paper by the noted optical experimenter H. E. Ives [8], whose 1938 experimental confirmation, in collaboration with Stilwell, of the quadratic Doppler effect lent support to the relativistic clock retardation.

The development of my thesis that a thorough grasp of the philosophical foundations of the special theory of relativity is indispensable to its illuminating historical study will turn out to yield an important consequence for the pedagogy of the theory. Namely, the clear recognition that the standard presentation of the theory in many physics texts such as is given by Richtmeyer, Kennard, and Lauritsen in their well-known *Introduction to Modern Physics* [9] must be revamped, because it inverts the *logical order* of Einstein's ideas and badly beclouds their epistemological anchorage.

As my first case in point, I turn to the principle of the constancy of the speed of light in the special theory of relativity – hereafter called ‘the light principle’ for brevity – which asserts that the speed of light is the same constant c in all inertial systems, independent of the relative velocity of the source and observer, and of direction, position, or time.

1. HISTORY AND PEDAGOGY OF THE LIGHT PRINCIPLE

The standard textbook presentation would have us believe that historically Einstein resolutely enunciated the light principle as a direct inductive generalization from the null outcome of the Michelson-Morley experiment. Thus we find, in the aforementioned well-known textbook on modern physics, that the statement of that experimental result is used in an attempt to confer credibility on the claim that “Einstein accordingly put the laws of the propagation of light in the forefront of the discussion.” And this statement is then immediately followed by the assertion that Einstein based his special theory of relativity on two postulates, the principle of relativity and the light principle.

I now endeavor to show that this kind of logico-historical account on which I was brought up is grossly misleading and unsound for the following reasons. (i) It asks the student to swallow a serious conceptual travesty, buttressed by a historical myth which explodes in the face of a clear comprehension of the philosophical foundations of the light principle. (ii) In *conceptual* respects, the standard textbook account in question is glaringly incompatible even with the text of Einstein's historically taciturn own fundamental 1905 paper on the special theory, a paper that by no means gives an adequate account, as we now know from Einstein

himself ([5], p. 228), of the full reasoning by which he groped his way to an espousal of the theory.

(1) It is logically indispensable that the statement of Einstein's light principle be *preceded* – as indeed it is in his 1905 paper – and not followed by the repudiation of the Newtonian conceptions of simultaneity, time interval, and distance in favor of the corresponding relativistic ones. For otherwise the student who is confronted with the assertion of the light principle will assign Newtonian meanings to its terms, a construal which will involve the assumption of the validity of the *Galilean* transformations. Hence in that context the assertion of the light principle will be a self-contradictory claim whose evident absurdity can hardly be removed by the invocation of the null-result of the Michelson-Morley experiment. And it is then also historically incredible that Einstein himself would have propounded and espoused so blatant an absurdity without *first* having supplanted the appropriate Newtonian conceptions by relativistic ones. Yet precisely this combination of logical and historical absurdities is foisted on the student by the rather standard textbook presentation of Richtmeyer, Kennard, and Lauritsen. After describing the null result of Michelson and Morley, these authors introduce the statement of the light principle to the Newtonian student by the following sentences:

Yet it seems quite impossible that light should move with the same velocity relative to each of two frames that are *moving relatively to each other*!

Einstein accordingly [sic] put the laws of the propagation of light in the forefront of the discussion

(2) But on turning to the pages of Einstein's fundamental paper of 1905, we find that it is not until the start of section 2 (§2) that he puts the light principle into the forefront of the discussion. And when he does give a statement of that principle at the beginning of the section, he includes a conceptually crucial qualification which is conspicuously absent from the aforementioned textbook account. Namely, he notes pointedly that the one-way velocity of light to which the principle refers is based on the definition of simultaneity which he had given in §1. In that *first* paragraph, he had given an epoch-making new philosophical treatment of the problem of simultaneity *within a single inertial system*. And not until §2 did he state the equivalent of the light principle and then show at the end that the latter entails discordant judgments of simultaneity as between Galilean frames moving relatively to one another.

Specifically, the conceptual innovation of Einstein's §1 is the anti-Newtonian view that the metrical simultaneity of two spatially separated events involves a convention *within any given inertial system*. It is thus a grievous philosophical blunder to suppose, as is done in many expositions, that the repudiation of Newton's conception of simultaneity by the theory occurs, *in the first instance*, in the context of the relative motion of different inertial systems [10]. And this philosophical misconstrual of Einstein's conception of simultaneity precludes awareness that his conception in §1 rests on two essential physical assumptions (to be stated below) *which must not be identified with the two well-known postulates of Einstein's §2*. Hence the philosophical misplacement of the *primary* locus of Einstein's conceptual departure from Newton in regard to simultaneity issues in the following serious penalty for the investigator of the history of the theory: the lack of philosophical mastery on the part of the historian will conceal from him that there exists the historical question as to the grounds on which Einstein felt entitled, in §1 of his 1905 paper, to make the particular two physical assumptions undergirding his doctrine of the conventionality *and* relativity of the simultaneity of spatially separated events! [11].

Einstein formulates this doctrine by stating in §1 that the equality of the one-way velocities of light in opposite directions within the same inertial system is a matter of *definition* rather than of prior physical fact: The facts of the *temporal order* allow the round-trip time of light for a given path connecting two points P_1 and P_2 to be split into any two parts as the respective one-way transit times in the two opposite directions P_1P_2 and P_2P_1 . And the two physical assumptions providing the logical underpinning of his innovative conception of simultaneity are *not* the two well-known postulates of §2 but the following two different *non-Newtonian* physical assumptions: (i) Within the class of physical events, material clocks of identical constitution do not define uniquely obtaining or 'absolute' relations of simultaneity under transport. (ii) Light is the fastest signal in a vacuum in the following ordinal sense: no kind of causal chain (moving particles, radiation) emitted in a vacuum at a given point A together with a light pulse can reach any other point B earlier, as judged by a local clock at B which merely orders events there in a metrically arbitrary fashion, than this light pulse.

Einstein sets forth his *definition* of metrical simultaneity without

stating these two assumptions, which are a (partial) set of premises for his conception of simultaneity as being both convention-laden and relative. But philosophical analysis discloses their crucial presence in the logical foundations of his doctrine of simultaneity as set forth in §1. Hence philosophical analysis prevents the historian from overlooking the point that Einstein must have made the two stated assumptions and must have justified them to himself before finally stating their upshot in very concise form. And in this way philosophical analysis makes for awareness that there is an historical question as to the grounds on which Einstein convinced himself of the two underlying assumptions when writing his 1905 paper.

That the philosophical discernment of the presence of the two tacitly made physical assumptions does indeed give a correct steer to the student of the genesis of the special theory of relativity is attested by the historical account that Einstein gave to Wertheimer as well as by his 'Autobiographical Notes'. Wertheimer writes ([5], p. 228, note 7):

I wish to report some characteristic remarks of Einstein himself. Before the discovery that the crucial point, the solution, lay in the concept of simultaneity, axioms played no role in the thought process of this Einstein is sure. (The very moment he saw the gap, and realized the relevance of simultaneity, he knew this to be the crucial point for the solution.)

More particularly, when talking to Wertheimer, Einstein made a decisive comment on the exposition of the theory given in his joint book with Infeld, where the two postulates of §2 of the 1905 paper are presented in standard textbook fashion as the fundamental axiomatic starting point of the theory [12]. He tells us ([5], p. 228, note 7) that this way of presenting the theory

is not at all the way things happened in the process of actual thinking. This was merely a later formulation of the subject matter, just a question of how the thing could afterwards best be written. The axioms express essentials in a condensed form. Once one has found such things one enjoys formulating them in that way.

And Einstein's 'Autobiographical Notes' record [13] that all attempts to develop the theory

were condemned to failure as long as the axiom of the absolute character of time, viz., of simultaneity, unrecognizedly was anchored in the unconscious. Clearly to recognize this axiom and its arbitrary character really implies already the solution of the problem. The type of critical reasoning which was required for the discovery of this central point was decisively furthered, in my case, especially by the reading of David Hume's and Ernst Mach's philosophical writings.

In addition to serving as a pathfinder for the historian of science, philosophical awareness of the foundations of Einstein's conception of simultaneity can provide prophylaxis against certain kinds of historical and pedagogical errors. I noted that philosophical analysis exhibits the *logical* dependence of Einstein's doctrine of simultaneity on the prior assumption that light is the fastest possible signal in a vacuum, an assumption to which I refer as 'the limiting assumption'. But there is a widespread failure to realize that the limiting assumption is thus presupposed both by the definition of simultaneity in Einstein's §1 and by the light principle in §2. This philosophical failure inspires the incorrect supposition that the limiting assumption still requires deductive justification within the theory by the time the laws of velocity addition are derived from the Lorentz transformations. The further *erroneous* belief that the limiting assumption is actually deducible from the velocity-addition laws -- which it is *not* [14] -- then begets the *historical* falsehood that these addition laws were the basis on which Einstein was first able to convince himself of the truth of the limiting assumption. Hence one boggles at the extent to which otherwise excellent physics books incorrectly present the limiting assumption as a deduction from the velocity-addition laws [15].

2. CONTRACTION AND TIME-DILATION HYPOTHESES

The charge that the aether-theoretic Lorentz-Fitzgerald contraction hypothesis and the aether-theoretic Lorentz-Larmor time-dilation hypothesis were severally and collectively *ad hoc* has figured prominently in the debate between the aether-theoretic conception of the Lorentz transformations and Einstein's rival relativistic interpretation of them. And the aether-theoretic Lorentz-Fitzgerald contraction hypothesis has come to be the most widely used textbook and classroom illustration of an *ad hoc* auxiliary hypothesis. Clarity as to the several senses in which an auxiliary hypothesis may be held to be *ad hoc* will permit determining in which of these several senses, if any, either or both of the aforementioned aether-theoretic auxiliary hypotheses are in fact *ad hoc*. Hence such clarity can contribute to *historical* understanding of the logical and psychological factors which enabled Einstein's relativistic conception of the Lorentz transformations to gain acceptance at the

expense of the earlier rival aether-theoretic interpretation of them.

Let me begin with an elucidation and appraisal of Einstein's own rejection of the Lorentz-Fitzgerald contraction hypothesis. It will be convenient to have a name for an outcome of a particular *kind* of experiment that is embarrassing to a previously successful theory *T*, as in the case of the embarrassment of the original aether theory by the null result of the Michelson-Morley experiment. Using the first letter of the word 'embarrassing' as a prefix, I name such an embarrassing result an *E*-result of the particular kind of experiment with respect to the theory *T*. It is to be well understood that in the context of any one particular theory, the specification of the attributes which characterize the given *kind* of experiment as such and distinguish it from other kinds is given by reference to the kind of conditions under which the theory *T* yields specified particular values of the variables ingredient in its postulates. This specification is required to give precise meaning to the question whether in the context of a given theory, an auxiliary hypothesis which explains an embarrassing result of *one kind* of experiment lends itself to independent test in at least one *other kind* of experiment.

Speaking of Einstein's reaction to the Lorentz-Fitzgerald contraction hypothesis, Wertheimer reports ([5], pp. 218-219) that

for Einstein the situation was no less troublesome than before — he felt the auxiliary hypothesis to be a hypothesis *ad hoc*, which did not go to the heart of the matter — He felt that the trouble went deeper than the contradiction between Michelson's actual and the expected result.

Since Einstein thus forsook the Lorentz-Fitzgerald version of the aether theory as *ad hoc* in his quest for a new rival theory, I deem it reasonable to take Einstein's rejection of the contraction hypothesis as '*ad hoc*' in this context to be tantamount to the following two-fold claim: (i) Prior to the *E*-result of the Michelson-Morley experiment, no other kind of experiment had an outcome providing support for the contraction hypothesis. (ii) It is to be expected that the Lorentz-Fitzgerald modification of the aether theory will *not* be confirmed by subsequent tests of a kind *different* from the Michelson-Morley experiment. And this conjecture serves as a reason for not accepting the Lorentz-Fitzgerald contraction hypothesis as an explanation of the null result of the Michelson-Morley experiment. Thus I take Einstein's rejection of the contraction hypothesis as *ad hoc* not only to allow but to assert that this

hypothesis is indeed 'independently testable', that is, testable by experiments of a kind different from the Michelson-Morley experiment. For I construe Einstein's pejorative usage of the term '*ad hoc*' to refer to the posited fact that though independently testable, the contraction hypothesis would fail to secure subsequent independent experimental confirmation *as against the claims of a new rival theory*.

We shall see after some analysis that if Einstein did consider the contraction hypothesis as *ad hoc* in *this* sense, he was quite right. The analysis that vindicates Einstein will also completely *refute* the standard textbook indictment of the contraction hypothesis, which charges this hypothesis with being *ad hoc* in the quite different sense of *not* being testable *independently* of the Michelson-Morley experiment. Let us appraise the claim that the Lorentz-Fitzgerald contraction hypothesis is *ad hoc* in this widely alleged sense. To do so, it is essential that we remedy the imprecision of the concept of independent testability encountered in the literature. The existence of such imprecision and the need for removing it emerge from the following two sets of considerations.

In the first place, reference is made to the set of *all* observational consequences of a given theory *T*, when it is claimed that the contraction hypothesis has no consequences by which it might be tested other than the observations in a Michelson-Morley type of experiment. But the set of all observational consequences of a given sophisticated theory *T* can have a well-defined membership only if we delimit and, if possible, specify the rules of correspondence (sometimes misleadingly called 'operational definitions') which, in conjunction with the postulates of *T*, are held to constitute the given theory and which anchor the postulates of *T* in the 'observational' base. For in the absence of some such delimitation or 'freezing' of *T* in a given stage of its development, the theoretical terms of the postulates of *T* represent 'open' concepts in the following sense: they admit the adjunction of further rules of correspondence ('operational definitions') to those constituting the merely partial empirical interpretation of the postulates of *T* at any given time. And this open-textured character of the theoretical terms of *T* would then make for a corresponding imprecision or open-endedness in the membership of the class of observational consequences of *T*. Hence the latter class must be relativized to a specific, delimited set of rules of correspondence.

In this way, the systemic attribute of independent testability of an auxiliary hypothesis and also that attribute's negation of being *ad hoc* become correspondingly relativized to a delimited set of rules of correspondence. In the context of this proviso of relativization, consider an auxiliary hypothesis *H* which is introduced into the framework of a given theory *T* in response to an experimental outcome which is an *E*-result with respect to *T*. It is clear that the possession by *H* of the systemic attribute of being independently testable and thus *not ad hoc* within the framework of *T* in no way depends on whether the advocates of *H* are *aware* of such independent testability. To suppose that there is any such dependence is just as erroneous as to maintain that whether a mathematical proposition is a theorem in a given axiom system depends on whether mathematicians possess the psychological attribute of realizing that the given proposition is indeed a theorem. We shall see that the contraction hypothesis indeed does *not* constitute an *ad hoc* modification of the aether theory in the sense now under discussion. For I shall demonstrate that its confirmation is logically possible in an experiment different from the Michelson-Morley type. Hence, if Lorentz and Fitzgerald were in fact unaware of and disbelieved in the latter independent testability of their auxiliary hypothesis, their unawareness and disbelief cannot possibly render that hypothesis systemically *ad hoc*. If these theoreticians did espouse their contraction hypothesis while mistakenly believing it to be systemically *ad hoc*, this espousal would merely establish their own methodological culpability in this respect. In that case, their espousal of the contraction hypothesis can be said to have been *psychologically ad hoc*. But the proof below of the independent testability of their contraction need take no cognizance of the beliefs which they actually entertained about its independent testability.

There is a second respect in which the concept of independent testability requires more precise statement than any given heretofore in the literature to my knowledge. The need for this refinement has been pointed out to me by Carl Hempel in private correspondence [16]. Hempel notes that strictly speaking, *no* auxiliary *H* which is offered to save a theory *T* from an *E*-result is independently testable by itself or is ever *ad hoc* by virtue of failure to be independently testable in isolation from *T*. For *H* entails testable observational consequences only in conjunction with one or more of the basic or derivative principles of *T*. Hence

Hempel points out that the concepts of independent testability and of being *ad hoc* must make due allowance for this contextual character of the observational import of *H*. And he suggests that this might be attempted along the following lines: Given that a theory *T* not containing *H* entails a false observational consequence *F*, then an auxiliary hypothesis *H* is systemically *ad hoc* in the context of *T*, if the modification of *T* under *H* — which we call *TH* — has the same observational consequences as *T* with the sole exception of the single observationally false *F*. Thus, if *H* is to be *ad hoc*, all observational consequences of *T* other than *F* must be identical with those of *TH*. And it is to be understood here that the observational consequence *F* covers an infinite class of observation statements which differ only in the places and times to which they pertain.

Now Hempel notes that this definition of '*ad hoc*' involves the concept of a single observational consequence pertaining only to the outcome of a particular kind of experiment. And he questions whether the latter concept can be circumscribed in purely logical terms so as to ever allow an auxiliary hypothesis *H* to qualify as *ad hoc* by meeting the requirements of this definition. For Hempel doubts that any *H* can ever qualify or fail to qualify as *ad hoc* on the strength of the feasibility of a specification in purely logical rather than 'denotative' terms of what constitutes (i) one particular or single kind of experiment, and (ii) one single observational consequence *F* of *T* pertaining to the outcome of this one particular kind of experiment. And his reasons for this doubt are as follows. First, any attempt to provide the required purely logical specification of the one kind of experiment and of the content of *F* would always have to exclude what are, in some imprecise intuitive sense, variants upon the one experiment, thereby also excluding the outcomes of these variant experiments from the range of occurrences covered by *F*. Yet in the case of these variants as well, *TH* would predict a different outcome from the one entailed by *T* alone, so that *TH* differs in its observational import from *T* not only with respect to *F*. Thus no *H* could qualify as *ad hoc* in the sense of the definition. And second, any proof that a particular *H* is not *ad hoc* might be contestable as inconclusive even though the proof adduces the existence of two kinds of experiment for each of which *TH* predicts a different outcome from the one yielded by *T* alone. For in the absence of a purely logical delimitation of what con-

stitutes a single kind of experiment, it might be claimed that the two different kinds of experiment are merely subvarieties of one single kind of experiment.

On the basis of these doubts that the attribute of being *ad hoc* can be defined in purely logical terms, Hempel suggests that the methodologically important notion of an *ad hoc* hypothesis involves the following idea, which is not specifiable in purely logical terms: an auxiliary *H* which enables a theory *T* to explain an *E*-result in conjunction with *H* is *ad hoc* if it does *not* have any observational consequences that are *significantly or interestingly different* from the *E*-result.

I regard as sound Hempel's doubt that the property of being a systematically *ad hoc* hypothesis can be defined in purely logical terms. And I concur with his claim that the specification of *F* is made denotatively, as it were, on the basis of judgments pertaining to the particular theoretical system at issue. Hence I invoke the concept of a *significantly different* observational consequence when now giving specific meaning to the question of whether the Lorentz-Fitzgerald hypothesis is *ad hoc* in the context of the aether theory. To do so, I must now specify what is to be understood by a Michelson-Morley type of experiment in contradistinction to other kinds and by the observational consequence of the aether theory pertaining to the outcome of this kind of experiment. The relevant observational consequence *F* of the aether theory is that the round-trip time T_v of light for the vertical arm of the interferometer, which is perpendicular to the direction of the earth's motion, is

$$T_v = \frac{2l}{(c^2 - v^2)^{1/2}}$$

while the round-trip time T_h for the horizontal arm, which points in the direction of the earth's motion, is [17]

$$T_h = \frac{2l}{(c^2 - v^2)^{1/2}} \cdot \frac{1}{(1 - \beta^2)^{1/2}},$$

where $\beta \equiv v/c$. And this observational consequence *F* pertains to *any* experiment which yields an observational comparison of the two round-trip speeds for the case of arms of *nearly equal* length *l* located in an inertial system having a constant velocity *v* through the aether. Any experiment furnishing this observational comparison is thus classifiable

as belonging to the one Michelson-Morley type of experiment to which the deduction F from the theory pertains. In the light of Hempel's cautions, the question whether the Lorentz-Fitzgerald hypothesis constitutes an *ad hoc* modification of the aether theory therefore takes the following form: Is the Lorentz-Fitzgerald hypothesis testable in any kind of experiment which differs significantly and interestingly from the specified Michelson-Morley type? By demonstrating now that the answer to this question is decidedly affirmative, I shall establish that the Lorentz-Fitzgerald hypothesis is *not* systemically *ad hoc*.

The Lorentz-Fitzgerald hypothesis asserts that the horizontal arm of the Michelson interferometer is of contracted length $l \cdot (1 - \beta^2)^{1/2}$ rather than of length l . And once the length l in the expression for the horizontal round-trip time T_h is replaced by this contracted length, T_h becomes *equal* to T_v , and the *difference* between them *vanishes*, as suggested by the null result of the Michelson-Morley experiment. That the Lorentz-Fitzgerald hypothesis has observational consequences whose *confirmation* is logically possible independently of the Michelson-Morley type of experiment can now be shown by demonstrating the following: The Lorentz-Fitzgerald modification of the aether theory yields *different* observational consequences from those entailed by the aether theory for the case of the type of experiment performed by Kennedy and Thorndike in 1932. Although the Kennedy-Thorndike kind of experiment also employs an interferometer, it differs importantly from the Michelson-Morley type. For, in the first place, *as measured by laboratory rods*, the horizontal and vertical arms of the Kennedy-Thorndike experiment are not at all equal but are made as different in length as possible, thereby differing observationally with respect to the values of the relevant theoretical variable of length. And, in the second place, unlike the apparatus in the Michelson-Morley experiment, the interferometer of the Kennedy-Thorndike experiment does not have a constant velocity v in the aether by remaining in a single inertial system: instead the apparatus of the Kennedy-Thorndike experiment acquires different values of the relevant theoretical variable of velocity by being transported to various inertial systems via the diurnal rotation and annual revolution of the earth. If the unequal lengths of the vertical and horizontal arms as measured by rods in the laboratory have the values L and l , respectively, then the *difference* between the vertical and horizontal

round-trip times of light entailed via the Lorentz-Fitzgerald hypothesis is *not* the same as the one entailed by the original aether theory. Specifically, in the case of the Kennedy-Thorndike experiment, the Lorentz-Fitzgerald hypothesis entails that the difference $T_r - T_h$ have the non-vanishing value

$$T_r - T_h = \frac{2(L - l)}{(c^2 - v^2)^{1/2}}$$

which *varies* with the diurnally and annually changing velocity v of the apparatus relative to the aether. Were it to materialize, the variation of this quantity would serve to detect any existing velocity v of the apparatus relative to the aether even on the assumption of a Lorentz-Fitzgerald contraction. But *without* the Lorentz-Fitzgerald hypothesis, the original aether theory yields the *different* non-vanishing, variable quantity

$$T_r - T_h = \frac{2}{(c^2 - v^2)^{1/2}} \left(L - \frac{l}{[1 - (v^2/c^2)]^{1/2}} \right).$$

It is evident that it is logically possible for a Kennedy-Thorndike type of experiment to *confirm* the quantitative predictions of the Lorentz-Fitzgerald hypothesis as against those of the original aether theory independently of the Michelson-Morley experiment. And this *logical* fact shows that the Lorentz-Fitzgerald hypothesis was *not ad hoc* in the systemic sense of Hempel's suggestion.

Furthermore, it is a matter of *empirical* fact that the Kennedy-Thorndike experiment of 1932 did *not* yield a shift in the interference fringes corresponding to the time difference deduced from the Lorentz-Fitzgerald hypothesis. In fact, just like the Michelson-Morley experiment, the Kennedy-Thorndike experiment had a negative outcome in the sense that there were no fringe shifts. Thus, one is entitled to claim that the Kennedy-Thorndike experiment failed to produce the kind of positive effect whose occurrence would have served to *confirm* the Lorentz-Fitzgerald hypothesis. But it would be an *error* to suppose that the *non-obtaining* of this particular kind of confirmation suffices to prove that the Lorentz-Fitzgerald hypothesis was *falsified* by the null result of the Kennedy-Thorndike experiment! For we shall now see that the

adjunction of the *further* auxiliary hypothesis of time dilation to the Lorentz-Fitzgerald hypothesis does enable the thus *doubly* amended aether theory to explain the null outcome of the Kennedy-Thorndike experiment while upholding the Lorentz-Fitzgerald hypothesis. And it will then become apparent that the justification for rejecting the Lorentz-Fitzgerald hypothesis along with the doubly amended aether theory depends on having philosophical reasons for accepting Einstein's rival theory of special relativity to the exclusion of the doubly amended aether theory. This finding *supersedes* pp. 392 and 394 of Chapter 12!

The velocity-dependent time-difference

$$\frac{2(L-l)}{(c^2 - v^2)^{1/2}}$$

yielded by the Lorentz-Fitzgerald hypothesis can be expressed alternatively as

$$T_r - T_h = \frac{2(L-l)}{c(1 - \beta^2)^{1/2}}.$$

Now suppose that in addition to the Lorentz-Fitzgerald hypothesis, one accepts the further Lorentz-Larmor auxiliary assumption that the rates of the clocks in a moving system are reduced by a factor of $(1 - \beta^2)^{1/2}$ as compared to the aether-system clocks. On this assumption of 'time dilation', the time-difference $T_r - T_h$ assumes the constant value

$$T_r - T_h = \frac{2(L-l)}{c}$$

which is independent of the velocity of the apparatus through the aether, in conformity to the null result of the Kennedy-Thorndike experiment. Thus, when amended by both the Lorentz-Fitzgerald hypothesis and the time-dilation, the aether theory does account for the actual outcome of the Kennedy-Thorndike experiment.

Moreover, purely mathematically the doubly amended variant of the aether theory permits the deduction of the Lorentz transformation equations no less than does Einstein's special theory of relativity. And this aether-theoretic deducibility of the Lorentz transformations now permits us to see that even the *conjunction* of the Lorentz-Fitzgerald

hypothesis with the assumption of the time dilation is not *ad hoc*. That the latter conjunction of auxiliary hypotheses is indeed testable in a kind of experiment which is independent of *both* the Michelson-Morley and Kennedy-Thorndike types is shown by the example of the so-called 'quadratic' optical Doppler effect as follows: Being *mathematically* identical with the space and time transformations of the special theory of relativity, the Lorentz transformations of the doubly amended aether theory entail an optical Doppler effect which is *quantitatively different* from the one that is deducible from the original aether theory [18]. Hence, the rejection of the doubly amended aether theory cannot be justified by claiming that the conjunction of its two auxiliary hypotheses is *ad hoc*, unless the term '*ad hoc*' is understood in a sense *different* from the one rendered by Hempel's suggestion.

I maintain that there is a very useful and interesting *different* sense of the term '*ad hoc*' which becomes relevant when the aether-theoretic conjunction of the Lorentz-Fitzgerald contraction and the time-dilation is appraised in the context of the conceptual rivalry between the doubly amended aether theory and the special theory of relativity. And that different sense of '*ad hoc*' is associated with a correspondingly different sense of 'independent testability' which is the following. Since the observational consequences of the *aether-theoretic interpretation* of the Lorentz transformations are the same as those of their rival relativistic interpretation, the aether-theoretic interpretation can have no observational consequences that are different from those of the rival special theory of relativity. Hence there can be no observational consequences which would support the doubly amended aether theory as against the new rival special theory of relativity, a theory that refuses to postulate the existence of some one preferred inertial aether frame when there is no kind of physical foundation for doing so. In the light of the absence of *this kind* of independent testability of the combined Lorentz-Fitzgerald and Lorentz-Larmor auxiliary hypotheses, their espousal solely for the sake of upholding the aether theory *to the exclusion* of the special theory of relativity can be said to be '*ad hoc*' in the new sense [19]. And this new sense clearly makes the *ad hoc* attribute of the conjunction of the two auxiliary hypotheses relative to *two* theories which are *conceptually* rivals of one another though not differing in observational import [20].

My analysis has endeavored to show that *prior* to the availability of the special theory of relativity, Lorentz, Fitzgerald, Larmor, and the remaining aether-theoreticians could not justly be accused of having put forward auxiliary hypotheses which were, in fact, systemically *ad hoc* in *either* of our two senses. On the strength of the philosophically mistaken assessment of these auxiliary hypotheses as either singly or collectively systemically *ad hoc* in the sense of Hempel's suggestion, the historical conclusion has been drawn that the original theorizing of the aether-theoreticians involved a grave infraction of scientific method of which they ought to have been aware. In this way, a philosophical error *unfoundedly* generated the historical allegation that the aether-theoreticians were unable to envision independent tests for their auxiliary hypotheses and hence were methodologically culpable from the beginning for espousing them nonetheless. And laboring under the philosophical misconception that the auxiliary hypotheses are systemically *ad hoc* in the sense of Hempel's suggestion could dissuade an historian of science from making the effort to uncover the kind of *historical* evidence which alone can show whether they were *psychologically ad hoc* either severally or collectively.

Let me conclude by noting that in a paper of 1951 entitled *Is There an Aether?*, P. A. M. Dirac attempted to resuscitate the aether with the aid of quantum mechanics [21]. That the aether may yet further engage the attention of philosophers of science in ways relevant to the concerns of the historian of science emerges from the following statements by Dirac, who wrote:

... We may set up a wave function which makes all values for the velocity of the aether equally probable. Such a wave function may well represent the perfect vacuum state in accordance with the principle of relativity.

... We thus see that the passage from the classical theory to the quantum theory makes drastic alterations in our ideas of symmetry. A thing which cannot be symmetrical in the classical model may very well be symmetrical after quantization. This provides a means of reconciling the disturbance of Lorentz symmetry in space-time produced by the existence of an aether with the principle of relativity.

I hope that this examination of the philosophy and history of the relativistic light principle and of the aether-theoretic auxiliary hypotheses has served to illustrate that philosophy is indeed instrumental in illuminating the *genesis* of the conceptual innovations wrought by a physical theory.

See Append.
§61

3. SUMMARY

There are several specific ways in which the philosophy of science is instrumental in illuminating the *genesis* of the conceptual innovations wrought by a particular physical theory. The unraveling of the history of Einstein's special theory of relativity is used to maintain concretely that the philosophy of science does have far-reaching relevance to the attainment of several particular cardinal objectives of the historian of science. The development of this thesis by reference to special relativity focuses on (i) the principle of the constancy of the speed of light and (ii) the evaluation of the charge that the Lorentz-Fitzgerald contraction hypothesis and the Lorentz-Larmor time-dilation hypothesis were severally and collectively *ad hoc* modifications of the aether theory of light propagation. The analysis yields a corollary for the pedagogy of the special theory of relativity: the standard textbook presentation of that theory must be revamped, for it inverts the logical order of Einstein's ideas and beclouds their epistemological anchorage.

BIBLIOGRAPHY AND NOTES

- * This chapter is the revised text of the author's vice-presidential address to the History and Philosophy of Science Section (I.) of the American Association for the Advancement of Science, delivered 29 December 1963 during its Cleveland meeting.
- [1] A. Grünbaum, *Philosophical Problems of Space and Time*, Knopf, New York, 1963.
 - [2] A. Grünbaum, 'The Relevance of Philosophy to the History of the Special Theory of Relativity', *J. Phil.* **59**, 561 (1962); 'The Special Theory of Relativity as a Case Study of the Importance of the Philosophy of Science for the History of Science', *Am. Mat.* **57**, 257 (1962) [reprinted in *Philosophy of Science* (ed. by B. Baumrin), Wiley, New York, 1963, vol. 2]; 'The Genesis of the Special Theory of Relativity', in: *Current Issues in the Philosophy of Science* (ed. by H. Feigl and G. Maxwell), Holt, Rinehart and Winston, New York, 1961, pp. 43-53.
 - [3] R. S. Shankland, *Am. J. Phys.* **31**, 47 (1963). The important second installment of this paper, published a decade later when this volume was in press, is discussed in the *Appendix* under 'Chapter 12'.
 - [4] This report constitutes chapter 10 of Max Wertheimer's *Productive Thinking* [5].
 - [5] M. Wertheimer, *Productive Thinking*, Harper, New York, 1959.
 - [6] R. S. Shankland, 'Albert A. Michelson at Case', *Am. J. Phys.* **17**, 487 (1949).
 - [7] R. S. Shankland, 'Conversations with Einstein', *Am. J. Phys.* **31** (1963), p. 55 and p. 49, note 9.
 - [8] H. E. Ives, 'Revisions of the Lorentz Transformations', *Proc. Am. Phil. Soc.* **95**, 125 (1951).
 - [9] F. K. Richtmeyer, E. H. Kennard, and T. Lauritsen, *Introduction to Modern Physics*, McGraw-Hill, New York, 1955, pp. 56-57.

- [10] For a specific refutation of this error, see my pp. 692–693 of Chapter 20, which amend pp. 342–370 of Chapter 12, as explained in the *Appendix* under 'Chapter 12'.
- [11] As shown by the putative 'quasi-Newtonian' universe of pp. 689–693 (Chapter 20), the relativity of simultaneity is *not* entailed by the latter's convention-ladenness.
- [12] A. Einstein and L. Infeld, *The Evolution of Physics*, Simon and Schuster, New York, 1942, p. 186.
- [13] See A. Einstein, 'Autobiographical Notes', in *Albert Einstein: Philosopher-Scientist*, (ed. by P. A. Schilpp), Library of Living Philosophers, Evanston, Ill., 1949, pp. 52–53.
- [14] For a proof that this belief is indeed erroneous, see Chapter 12, pp. 375–376.
- [15] For examples of the commission of this fallacy, see E. T. Whittaker, *A History of the Theories of Aether and Electricity*, Nelson, London, 1953, p. 38; and R. C. Tolman, *Relativity, Thermodynamics and Cosmology*, Oxford Univ. Press, Oxford, 1934, p. 26.
- [16] I thank Professor Hempel for his kind permission to cite this private correspondence and also for helpful comments on an earlier version of this chapter.
- [17] A *slight* inequality of the horizontal and vertical arms of the interferometer is required for the production of *neat* interference fringes. But I ignore this slight length discrepancy in my generally oversimplified presentation by using the same length l in the expressions for T_i and T_h : A rigorous detailed justification of this use of the same length l as part of my simplification is given in L. Silberstein, *The Theory of Relativity*, Macmillan, London, 1914, p. 76. On pp. 72–79, Silberstein gives an especially valuable account of the Michelson-Morley experiment and of the aether-theoretic contraction hypothesis. See also J. Aharoni, *The Special Theory of Relativity*, Oxford University Press, Oxford, 1959, pp. 270–273.
- [18] See Section 7 of Einstein's fundamental paper of 1905 on the special theory of relativity, and M. von Laue, *Die Relativitätstheorie*, Vieweg, Braunschweig, 1952, vol. 1, p. 20.
- [19] This account of the sense in which the two combined auxiliary hypotheses can be said to be *ad hoc* supersedes the treatment in Chapter 12, p. 392.
- [20] It is an open question whether this new sense of '*ad hoc*' can be explicitly defined quite generally by reference to any two rival theories of an appropriate kind, and to what extent that putative new general sense would have relevance to diverse episodes in the history of science.
- [21] P. A. M. Dirac, 'Is There an Aether?', *Nature* **168**, 906 (1951).

GENERAL RELATIVITY, GEOMETRODYNAMICS AND ONTOLOGY

I. INTRODUCTION

For nearly two decades before 1972, Professor John Wheeler pursued a research program in physics that was predicated on a monistic ontology which W. K. Clifford had envisioned in 1870 and which Wheeler (1962b, p. 225) epitomized in the following words: "There is nothing in the world except empty curved space. Matter, charge, electromagnetism, and other fields are only manifestations of the bending of space. *Physics is geometry.*" In an address to a 1960 Philosophy Congress (Wheeler, 1962a), he began with a qualitative synopsis of the protean role of curvature in endowing the one presumed ultimate substance, empty curved space, with a sufficient plurality of attributes to account for the observed diversity of the world. Said he (1962a):

... Is space-time only an arena within which fields and particles move about as 'physical' and 'foreign' entities? Or is the four-dimensional continuum all there is? Is curved empty geometry a kind of magic building material out of which everything in the physical world is made: (1) slow curvature in one region of space describes a gravitational field; (2) a rippled geometry with a different type of curvature somewhere else describes an electromagnetic field; (3) a knotted-up region of high curvature describes a concentration of charge and mass-energy that moves like a particle? Are fields and particles foreign entities immersed *in* geometry, or are they nothing *but* geometry?

It would be difficult to name any issue more central to the plan of physics than this: whether space-time is only an arena, or whether it is everything (p. 361).

For nineteen years, Wheeler and his co-workers, such as Charles Misner, developed some of the detailed physics of Clifford's ontology of curved empty space-time as an outgrowth of general relativity under the name of 'geometrodynamics' ('GMD'). In Wheeler's parlance, 'a geometrodynamical universe' is "a world whose properties are described by geometry, and a geometry whose curvature changes with time — a *dynamical* geometry" (Wheeler, 1962a, p. 361). But in a lecture at a 1972 conference,¹

¹ Conference on Gravitation and Quantization, held in October and November, 1972 at the Boston University Institute of Relativity Studies, directed by John Stachel. I am grateful to Professor Stachel for having given me the opportunity to attend this conference.

Wheeler disavowed his erstwhile long quest for a reduction of all of physics to space-time geometry.² In a brief notice of that Conference (*Nature* **240** [Dec. 15, 1972], p. 382), the pertinent part of this lecture was summarized as follows: "He [Wheeler] also developed the theme that the structure of space-time could only be understood in terms of the structure of elementary particles rather than the converse statement which he has advocated for many years."

To emphasize his new conception of the fundamental and indispensable ontological role played by entities or processes *other than* space-time, Wheeler repeatedly spoke of PRE-geometry. As I understood him, he sought to emphasize in this way that he now regards space-time not as the basic stuff of a monistic ontology but rather as an *abstraction* from the events in which quantum processes are implicated. That is to say, according to Wheeler's new conception of PRE-geometry, space-time is an abstraction from the constitution of physical events ONTOLOGICALLY no less than epistemologically! In Wheeler's *erstwhile* view of space-time as the only autonomous substance, space-time is absolute in the older familiar sense of being *empty*. In the perspective of Wheeler's new program of a PRE-geometric ontology, his earlier all-out geometric absolutism appears as a kind of ontological chauvinism. Wheeler now wishes to supplant that absolutism by a neo-Leibnizian view which is RELATIONAL in the sense that space-time structure is only one aspect of a quantum universe whose ontological furniture cannot be constituted out of space-time. I believe that one reason for Wheeler's explicit mention of Leibniz was to emphasize the relational character of his notion of pre-geometry.

In the minds of great scientists like Wheeler, there often is a subtle interplay between empirical and conceptual or philosophical promptings for abandoning no less than for espousing a major theory with its research program. It seems to me that this state of affairs need not at all betoken a gratuitous and pernicious apriorism: At least generally, there is no sharp divide between legitimate empirical and conceptual or philosophical reasons for the rejection (acceptance) of a major theory. Thus, I believe that predominantly philosophical considerations are likewise germane to the appraisal of certain facets of Wheeler's erstwhile purely geometrical ontology vis-à-vis the rival ontology adumbrated in his more recent

² This disavowal can now be seen to have been heralded by Wheeler's January, 1971 Foreword in Graves (1971, p. viii).

notion of pre-geometry. Hereafter, when I speak of 'geometrodynamics' or use its acronym 'GMD', I shall disregard Wheeler's basic change of view and will use this term to refer to his earlier relativistic theory of *empty* space, rather than to the ontologically more noncommittal standard version of general relativity. Even though it was *not*, of course, relativistic, Clifford's so-called 'Space-Theory of Matter' shared the essential ontological assumptions of Wheeler's GMD. Hence I shall also occasionally denote Clifford's theory-sketch by 'GMD'.

Partly in response to a 1972 GMD Symposium (*Journal of Philosophy* 69, No. 19 [October 26, 1972]) which focused on Graves (1971), the present chapter is partly devoted to an analysis and appraisal of some of the major facets of the ontology of GMD as developed by Clifford and Wheeler.

Before entering into such analysis and appraisal, we shall first be concerned in the next section with foundational aspects of the Riemannian metric of space-time which are central to Einstein's *standard* 1915–1916 general relativity theory, and with which I have not dealt before. Only thereafter shall we be ready to address ourselves to the ontological articulation of the GMD thesis that the metric geometry of the space-time manifold in which we live is the physical world's only autonomous substance. See Appendix §§62, 63

2. THE PHILOSOPHICAL STATUS OF THE METRIC OF SPACE-TIME IN THE GENERAL THEORY OF RELATIVITY

Three major alternative methods for physically grounding and theoretically circumscribing the metric of space-time up to a constant scale factor merit consideration, although none of them is wholly unproblematic. We shall consider two of them rather briefly before giving our main attention to a third, which has been adduced as foundationally most germane to the ontological vision of GMD. Let us begin with the most traditional of the three alternative methods.

(i) *The Method of Rods-and-Clocks*

Both in his 1917 Prussian Academy Lecture 'Geometry and Experience', and in his 'Autobiographical Notes', Einstein (1949, pp. 59–60) appealed to special rods and clocks in order to give a provisional explication of the physical significance of the space-time metric. He regarded this theoretical

foundation for the metric as provisional ontologically, because he noted (Einstein, 1949, p. 59) that macroscopic rods and clocks are 'objects consisting of moving atomic configurations' and in that sense not, as it were, 'theoretically self-sufficient entities'. In this same vein, others have gone on to point out that there is the following kind of serious question of principle concerning the ability of solid rods to fulfill the role of a metric standard in the general theory of relativity (hereafter 'GTR'): If such a rod is allowed to fall freely in a permanent gravitational field such as the earth's, its end-points will be either driven closer together or further apart by that field. Thus, *as judged by the space-time metric prescribed by the GTR*, the length of the rod will change, although the amount of that change will depend on the constitution of the rod. Making relativistic corrections for such deviations from rigidity may be very difficult, and may involve much of the *whole* GTR on pain of logical circularity.

But, as I emphasized elsewhere (Grünbaum, 1968, p. 367), the *ontology* of the GTR-metric which I espouse in Section 3(a) below *allows* fully that the following statement by Putnam (1963, p. 206) be true: "... the metric is implicitly specified by the whole system of physical and geometrical laws and 'correspondence rules.'"

Having seen some of the difficulties of reliance on solid rods even infinitesimally, one might follow Synge (1956, Ch. I, §§9, 13 and 14; 1960, Ch. III, §4) and use another method which is *chronometric*.

(ii) *Synge's Chronometric Method*

Synge writes (1956):

To measure time one must use a *clock*, a mechanism of some sort in which a certain process is repeated over and over again under the same conditions, as far as possible. The mechanism may be a pendulum, a balance wheel with a spring, an electric circuit, or some other oscillating system, and out of these one passes by idealisation to the concept of a *standard clock*...

Let us however make the concept of a standard clock more definite by thinking of it as an atom of some specified element emitting a certain specified spectral line, the 'ticks' of the clock corresponding to the emission of the crests of successive waves of radiation from the atom. By associating numbers 1, 2, 3, ... with these ticks, we have a scale for assigning times to events occurring in the history of the atom, or (since that would give us a very small unit of time) we may more conveniently associate with the ticks the numbers $a, 2a, 3a, \dots$, where a is some chosen small number, to be used universally for all clocks.

To base the measurement of time, as above, on a standard atomic frequency is very like the plan of basing the measurement of length on a standard wave length. But in relativity the concept of length is not an easy one, and it seems best to start with an atomic clock as the basic concept and introduce the idea of wave length at a later stage (pp. 14 and 15).

Synge goes on to claim (1956, pp. 15–16) that (a) For two neighboring events on the world line of a standard (atomic) clock whose coordinates are x^i and $x^i + dx^i$ respectively, the invariant infinitesimal *space-time* separation ds is equal to the *time* interval between them as measured by that clock, and (b) The chronometrically ascertained values ds for such pairs of events are equal to those of some function $f(x^i, dx^i)$ which is positive homogeneous of first degree in the coordinate differentials dx^i . This means that if the differentials dx^i are each multiplied by the same positive factor k , ds will be multiplied by that factor, so that

$$f(x^i, k \cdot dx^i) = k \cdot f(x^i, dx^i) \quad \text{for } k > 0.$$

An invariant line element $ds = f(x^i, dx^i)$ which has this mathematical property is said to be the metric of a *Finsler space* (Synge, 1956, p. 19).

Synge does not explain on what grounds it is assumed that the ds values which are ascertained chronometrically as specified are those of a Finslerian metric, but merely points out (1956, p. 19) that a Riemannian metric is only a special case of a Finslerian metric. For example, in a Finslerian space-time, the separation ds might have the *non*-Riemannian form

$$ds = (g_{mnrs} dx^m dx^n dx^r dx^s)^{1/4}.$$

Furthermore, the reason given by Synge (1956, p. 19) for now choosing the Riemannian species

$$ds = |g_{mn} dx^m dx^n|^{1/2}$$

of Finslerian line element is that this choice is dictated by the objective of formulating Einstein's theory. For the latter purpose, Synge selects in particular the kind of indefinite Riemannian metric which is of signature +2.³

Note that Synge renounced the foundational use of the solid rods of method (i) above in his chronometric ontology of the GTR space-time metric. By so doing, Synge's method forfeited such justification as the presumably *Pythagorean* infinitesimal metric behavior of rigid rods can

³ For a definition of the 'signature' of a quadratic differential form, see, for example, Weatherburn (1957, pp. 11 and 14).

furnish for adopting the Riemannian species of space-time metric within the Finslerian genus. Also, despite the essential immunity of the rate of atomic clocks to outside perturbations, it is not entirely clear that the observational use of atomic clocks – say, by tuning the vibrations of caesium atoms to synchronism with an oscillating electric circuit – can altogether avoid being epistemically parasitic on *non-relativistic* theory. Thus, it is not fully clear whether Synge's chronometric method is any more free from such possible epistemic parasitism than method (i) when the latter rods-and-clocks method attempts to correct solid rods for gravitational distortions.

So far, Synge's prescriptions for the chronometric determination of ds were confined to event pairs whose separations were time-like. But for any given event P of a relativistic space-time, there will be other events Q whose separation from P is space-like, and still other events R which can belong to the career of a free photon whose world-line also crosses P . Therefore, Synge proceeds to point out very interestingly (1956, p. 23) that if we are given the coordinate differences dx^i between P and any other nearby event as needed, then his chronometric method permits the *indirect* determination of the ds values even for event pairs (P, Q) whose separation is *space-like*! For given the values of dx^i corresponding to ten events which lie respectively on as many different world lines of atomic clocks that cross P , the chronometric determination of the ten respective separations of these events from P enables us to use ten corresponding equations

$$ds^2 = |g_{mn} dx^m dx^n|$$

to *calculate* the generally ten independent values g_{mn} at P . Although these g_{mn} values were obtained solely from determinations pertaining to the *time-like* world lines of ten atomic clocks, they qualify as the metric coefficients at P for *any kind* of infinitesimal space-time interval of which P is an end-point. Hence their substitution in the equation for the Riemannian metric of space-time now permits us to calculate, in turn, the *space-like* separation of a nearby event Q from P for which the coordinate differentials dx^i are known.

In Synge's construction, it can then be postulated that for the now known values of g_{mn} at P , the equation for the Riemannian metric will yield the null value $ds=0$ for the separation between P and nearby events R which can be linked to P by the world-lines of free photons and whose

coordinate differences dx^i from P are known (Synge, 1956, p. 23). Given these results, Synge (1956, Ch. I, §14) is able to furnish a second kind of ideal experiment for the measurement of a *space-like* separation, which is more direct than the first kind presented above and uses as apparatus only a standard clock and photons. In summary, he writes (Synge, 1956):

... we shall physicise the geometrical element d of separation between two adjacent events, P and Q .

First we find out by trial whether it is possible to make a material particle contain these two events in its history. If it is possible, then d is measured by the difference in the readings at P and Q of a standard clock carried by the particle which includes them in its history.

If it is impossible for a material particle to include both the events in its history, we know that PQ either lies on the null cone or is spacelike. We test whether it lies on the null cone by emitting photons from P in all the space-time directions permitted to photons. If one of these photons includes Q in its history, then PQ is null and we have $ds = 0$.

Suppose now that neither material particle nor photon can include both P and Q in its history. Then the space-time displacement PQ is spacelike. (p. 24)

But doubts have been raised concerning Synge's chronometric approach in an illuminating paper on the foundations of the space-time metric in the GTR by Ehlers *et al.* (1972). These three authors offer an important alternative to Synge's chronometric approach which will concern us below in the context of GMD. Hence it will be useful to quote *in extenso* from their paper (1972). Speaking of Synge's method, they write:

This procedure has two advantages. First, it uses as primitive a physical quantity that can, in fact, be measured locally and with extreme precision, and, secondly, it introduces as the primary geometric structure the metric, from which all the other structures can be obtained in a straightforward manner.

If the aim is a *deduction* of the theory from a few axioms, the *chronometric* approach is indeed very economical. If, however, one wishes to give a *constructive* set of axioms for relativistic space-time geometries, which is to exhibit as clearly as possible the physical reasons for adopting a particular structure and which indicates alternatives, then the chronometric approach does not seem to be particularly suitable, for the following three reasons. It seems difficult to derive *from the behaviour of clocks alone*, without the use of light signals, the *Riemannian* form for the separation,

$$ds^2 = |g_{ij} dx^i dx^j|^{1/2} \quad (1)$$

rather than some other, first-degree homogeneous, functional form in the dx^i (as, for instance, the Newtonian form $ds = g_i dx^i$). Postulating this form axiomatically, one foregoes the possibility of understanding the reason for its validity. The second difficulty is that if the g_{ij} are defined by means of the chronometric hypothesis, it seems not at all compelling

if we disregard our knowledge of the full theory and try to construct it from scratch – that these chronometric coefficients should determine the [geodesic] behaviour of freely falling particles and light rays, too. Thus the geodesic hypotheses, which are introduced as additional axioms in the chronometric approach, are hardly intelligible; they fall from heaven like Equation (1). Finally, once the geodesic hypotheses have been accepted, it is possible, in the theories of both special and general relativity, to construct clocks by means

of freely falling particles and light rays, as shown by Marzke [(1959)] and, differently, by Kundt and Hoffman [(1962)]. Thus, these hypotheses alone already imply a physical interpretation of the metric in terms of time. The chronometric axiom then appears either as redundant or, if the term 'clock' is interpreted as 'atomic clock', as a link between macroscopic gravitation theory and atomic physics: it claims the equality of gravitational and atomic time. It may be better to test this equality experimentally or to derive it eventually from a theory that embraces both gravitational and atomic phenomena, rather than to postulate it as an axiom.

For these reasons, we reject clocks as basic tools for setting up the space-time geometry [See also (Marzke and Wheeler, 1964)] and propose to use light rays and freely falling particles instead. We wish to show how the full space-time geometry can be synthesized from a few assumptions about light propagation and free fall (pp. 64–65).

Ehlers *et al.* no more explain than Synge himself on what grounds it is assumed that the chronometrically ascertained ds values are those of a Finslerian metric. But they object that in Synge's account the *speciation* of a physically unproblematic Finslerian space-time metric into a uniquely Riemannian one is gratuitous: Like manna, it 'falls from heaven'. It will therefore be of interest to note below whether their own alternative ontology of the GTR space-time metric can avoid a counterpart to Synge's 'manna' in its logical edifice. In any case, as I have already pointed out, by eschewing the solid rods of the more traditional rods-and-clocks method (i), Synge's chronometric method forfeits such justification as the presumably *Pythagorean* infinitesimal metric behavior of rigid rods can furnish for adopting the Riemannian species of space-time metric within the Finslerian genus.

(iii) *The Geodesic Method*

As was noted in Ch. 3, Section B, for surfaces in Euclidean 3-space, the prescription that a certain family of lines qualify as geodesics of the surface does *not* determine the metric of the surface up to a constant positive scale factor or even the Gaussian curvature of the surface modulo such a factor. For example, we saw there that on an ordinary table top, the familiar straight lines qualify as geodesics with respect to various metrics which can differ *other than* by a scale factor and which generate not only *different* partitions into equivalence classes of congruent intervals, but also *different* partitions of the class of *angles* into metrical equivalence classes.⁴ More

⁴ For further details, see Willmore (1959, pp. 87–88). The geodesic structure is often called the 'projective structure', because the term 'projective differential geometry' denotes the study of those properties of the geodesic paths themselves which are independent of any and all arc lengths defined on them.

generally, consider the case of the *positive definite* Riemann metrics (Weatherburn, 1957, pp. 12–14), familiarly encountered in the geometries of 3-dimensional physical spaces and discussed in Ch. 16, p. 472. The mere specification of those curves in a given space which are to count as geodesics of *such* a Riemannian metric does not determine a partition of the *angles* into metrical equivalence classes *and* does not suffice to single out a metric up to a constant positive scale factor k . The reason is that a mere *geodesic mapping* of the space onto itself does *not* determine the metric tensor g_{ik} even up to a conformal transformation (as defined on p. 19 of Ch. 1), let alone up to a similarity transformation or modulo k . On the other hand, if we prescribe both the partition of angles into metrical equivalence classes – via a suitable conformal mapping of the surface onto itself – *and* what paths are to be geodesics, then the metric (tensor) is prescribed to within k . In short, the combination of the conformal and geodesic (or ‘projective’) structures does circumscribe the metric (tensor) modulo k . ‘Projective’ is here understood in the sense of fn. 4, p. 735.

But the metrics of the space-times of the GTR are *indefinite* Riemannian metrics (cf. Ch. 16, p. 472 and Weatherburn, 1957, p. 12) of Minkowskian signature. As we shall see shortly, in the context of the GTR’s indefinite Riemannian space-time metric, Weyl (1922, pp. 179, 228–229, 313–314; 1923b, pp. 228–229; 1923a, pp. 18–19; 1949, p. 103) was able to impose two compatible requirements which generated conformal and geodesic (‘projective’) space-time structures respectively. And in this way, he effected a new theoretical circumscription of the GTR’s space-time metric modulo k .

We saw in our quotation from Ehlers *et al.* (1972) that these authors are concerned to provide a rationale for the Minkowskian kind of indefinite space-time *Riemann* metric used in the GTR by deriving it from assumptions and requirements which they regard to be more basic, as it were. By contrast, Weyl assumes at the outset, without any more rudimentary motivation, that in any GTR universe, the space-time metric (hereafter ‘ST-metric’) has this particular character.⁵ In essence, Weyl then goes on to obtain a *compatible* trio of assumptions by imposing the following two further conditions as to how the ST-metric is to be chosen: first, at each world point, light rays and only light rays are to be metrically *null* space-time trajectories, and second, the physical space-time trajectories

⁵ For a more detailed statement by Ehlers *et al.* of the divergences of their construction from Weyl’s, see their (1972, p. 68 and 68n).

of any and all freely falling mass particles are to count geometrically as time-like *geodesics* of the indefinite Riemannian ST-metric. The *compatibility* of these two conditions rests on the GTR assumption that the totality of free-fall mass particle trajectories passing through an event determines the light cone as its space-time boundary, since a free particle of positive rest mass, though always slower than light, can pursue a light signal arbitrarily closely (Weyl, 1922, pp. 228–229; Ehlers *et al.*, 1972, pp. 65–68). Let us see what contribution is made by each of these two compatible requirements to singling out, for any GTR world of given mass-energy distribution, a ST-metric which is unique modulo k and whose metric tensor has a *non-zero* determinant.

(1) *The requirement of the metrical nullity of light trajectories in space-time ('ST')*. Let us formulate more precisely Weyl's GTR requirement (Weyl, 1922, p. 228; 1923b, p. 228) that the ST trajectories of light pulses are coextensive with metrically null ST lines. At any given world-point P , consider the class D of ST directions of any and all infinitesimally neighboring distinct events Q , whose respective relative coordinates are given by a set dx^i in each case. Then Weyl first requires the metric tensor g_{ik} to be so chosen that within the class D of directions dx^i at any given world-point, all and only those ST directions which are physically permitted to light signals (either 'incoming' or 'outgoing') satisfy an equation of the form $ds^2 = g_{ik} dx^i dx^k = 0$. Amid letting the term 'photon' function interchangeably with the pre-quantum theoretic term 'light ray' in this context, let us designate this metrical nullity requirement for photons by the abbreviation 'photon- $ds=0$ '.

Thus, Weyl first requires g_{ik} to be so chosen that the particular directions which are singled out by infinitesimal light propagation at any given point P each be metrically null vectors. Since the infinitesimal light rays at P combine into a 'double' conical hypersurface in space-time, the latter's two lobes are thus *null* hypercones. By coupling the stated nullity stipulation as to the permissible kind of metric tensor with the GTR assumption concerning the behavior of photons at any one world-point, Weyl assures that infinitesimal light propagation determines an infinitesimal 'double' *null* cone at each point. Since photons have a vanishing rest mass, we shall also speak of *massless* (test) particles as generators of null cones. (But cf. p. 453.)

To what extent does Weyl's first requirement photon- $ds=0$ restrict the allowable metric tensor of the ST-metric? Let us elaborate on Weyl's own quite cryptic statement of the reasoning which issues in his well-known answer to this question. At any given ST point P , the equation $g_{ik} dx^i dx^k = 0$ generally contains ten independent values g_{ik} , i.e., there are at most ten such independent values. And at any point P , the nine sets of relative coordinates dx^i for nine nearby events respectively located on as many different photon trajectories through P furnish nine independent equations $g_{ik} dx^i dx^k = 0$ which suffice to determine the nine independent *ratios* of the ten independent values g_{ik} . It will be noted that the nine *null* directions of as many photons determine only the *ratios* of the g_{ik} values at P , because $ds=0$ for each of these directions. Furthermore, the ST directions dx^i allowed to photons at P are restricted to a double conical hypersurface. Hence it turns out that taking additional *null* directions would yield equations which are merely redundant with those corresponding to the initial nine null directions and would *not* restrict the g_{ik} values any further than their nine independent ratios do. Thus Weyl's first requirement photon- $ds=0$ determines only the nine independent *ratios* of the g_{ik} at any given ST point P . In particular the fixation of these ratios does *not* suffice to determine the ten independent *values* g_{ik} at P up to a positive factor k of proportionality which is the same constant from point to point. Contrast this with the feasibility of the determination of the ten individual *values* g_{ik} at P , at least up to a positive scale factor k which is the same constant from point to point, in Synge's aforementioned case of ten positive values ds for ten known *time-like* directions dx^i (see p. 733).

Since Weyl's first requirement photon- $ds=0$ determines *only* the *ratios* of the g_{ik} at any given point, this requirement *allows*, in any *one* coordinate system, alternative functions g_{ik} as follows: These differ *other than* by a positive constant scale factor k that is the same from point to point. Hence instead of determining the metric tensor or the metric ds up to a positive *constant* scale factor k , Weyl's first requirement determines the metric ds only up to *conformal* transformations of the metric tensor g_{ik} . The invariance of $ds=0$ for photon trajectories under conformal transformations means the of the metric tensor and only under such transformations means the following: If g_{ab} is a metric tensor which assures the metrical nullity $ds=0$ of photon ST trajectories, then this same nullity will be assured by just those non-zero metric tensors g_{ab}^* which are related to g_{ab} by a

positive (>0) multiplying function $f(x^i)$ of the coordinates. And clearly the case in which the conformal factor $f(x^i)$ is a positive *constant* k is only a very special case. Thus, Weyl's requirement $\text{photon-ds}=0$ defines a conformal structure for space-time and thereby allows an infinite class of *non-trivially* different metric tensors. Therefore, if, in this context, the allowed metric tensors are to be further restricted up to a positive constant scale factor k , one must demand more than the coextensiveness of the ST trajectories of light pulses with metrically null world-lines. And of course, the further restriction to be imposed must be compatible with Weyl's first requirement.

(2) *The requirement that the physical ST trajectories of freely falling mass particles qualify geometrically as time-like geodesics of the indefinite Riemannian ST metric.* Let it be granted for now that all freely falling (test) particles of *positive* rest mass, i.e., 'massive' (test) particles, have determinate ST trajectories, and that there are at least two such world-lines through every event (Weyl, 1968 reprint, p. 196; 1922, Appendix I, p. 314). Then Weyl compatibly supplements the conformal structure with a so-called '*projective*' structure on space-time by imposing the following further requirement (Weyl, 1923a, pp. 18–19; 1949, p. 103): The metric tensor g_{ik} of the indefinite Riemannian ST-metric ds of Minkowskian signature is to be *so chosen* that all ST trajectories of freely falling massive (test) particles are turned into time-like GEODESICS. In other words, Weyl also stipulates that the metric ds be chosen so as to enable the ST trajectories of massive particles to qualify or count as time-like geodesics via the equation $\delta \int ds = 0$.

Having also imposed this important *geodesicity* requirement, Weyl is able to show that the class of conformally related metric tensors singled out by the prior requirement $\text{photon-ds}=0$ is restricted further to a particular proper subclass whose members are pairwise related by any constant positive factor k : The factor k is constant in the sense of being the same from point to point rather than varying with the coordinates. One often speaks of any particular choice of this scale factor k as a trivial matter on the presumed ground that the laws of nature are invariant under all choices of a metrical unit of 'length'. But Weyl (1949, p. 83) remarks that in a quantum world, the laws of atomism restrict the allowable values of k to some extent. Regardless of whether k is thus restricted,

however, Weyl's construction – hereafter 'Weyl's geodesic method' – circumscribes the metric (tensor) modulo k . And since all of the other geometric structures can be obtained from the metric (tensor) in a straightforward manner, Weyl's geodesic method generates a unique metric geometry for the space-time of any GTR world of given mass-energy distribution.

As Weyl has been concerned to stress (Weyl, 1923a, p. 19; 1949, p. 103), if his geodesic method is not circular logically or epistemically – about which more below – then it determines the ST-metric modulo k both logically and epistemically, at least in principle, 'without reliance on clocks and rigid rods' (1923a, p. 19). By thus dispensing with the atomic clocks of Synge's method and with the rigid rods of Einstein's, Weyl has offered an important alternative foundation for the ST-metric of the GTR.

As we saw, the construction of the GTR ST-metric by Ehlers *et al.* (1972) – hereafter 'E & P & S' – differs from Weyl's at least as follows: The former three authors set themselves the task of deriving or providing a rationale for the Minkowskian kind of indefinite Riemann ST-metric, rather than merely circumscribing it to within k after it first 'fell from heaven', as they put it. But in carrying out their task, E & P & S define compatible conformal and projective structures on space-time by appealing, just as Weyl does, to the behavior of both massless and massive test particles, as postulated by the GTR. And after having defined the latter compatible structures, E & P & S specify additional necessary and sufficient conditions which then yield a Riemannian metric that is unique modulo k . In this way, they aim to improve on Weyl's version of the geodesic method by showing that the existence of a unique Riemannian metric structure need not be postulated beforehand in order to generate it by means of the joint behavior of massless and massive test particles.

Let us see somewhat more specifically how the E & P & S version of the geodesic method proposes to obtain by 'honest toil' – to use Russell's parlance – what Weyl obtained by 'theft', as it were. E & P & S (1972) write:

Whereas the ideas that light propagation determines a conformal structure $\%$ and free fall defines a projective structure $\mathcal{P}$ on space-time have been clearly spelled out by Weyl [(1921, 1923a and 1923b)], neither he himself nor anybody else, as far as we know, has used these two structures and their compatibility as fundamental, and derived the existence and uniqueness of an affine connection from these data. Rather, although Weyl emphasized repeatedly the fundamental roles of the structures $\%$ and $\mathcal{P}$ from a physical point of view,

in geometry he took the affine and metric structures as basic and considered the projective and conformal structures as arising from these by abstraction only, although also geometrically $\mathcal{C}$ and $\mathcal{P}$ are the more primitive (less restrictive) structures (p. 68).

This statement is the conclusion of the following summary they give of their more detailed reasoning:

(a) The propagation of light determines at each point of space-time the infinitesimal null cone and thus gives it a *conformal structure* $\mathcal{C}$. With respect to this, one can distinguish between time-like, space-like, and null directions, vectors, and curves, respectively, and one can single out as null geodesics those null curves contained in a null hypersurface (p. 65). These null geodesics will be shown to represent light rays.

(b) The motions of freely falling particles determine a family of preferred $\mathcal{C}$ -time-like curves, and by assuming this family to satisfy a generalized law of inertia, we show that free fall defines a *projective structure* $\mathcal{P}$ in space-time such that the world lines of freely falling particles are the $\mathcal{C}$ -time-like geodesics of $\mathcal{P}$.

(c) The conformal and projective structures thus defined are intimately related, as experience indicates: an ordinary particle (i.e. one with positive rest mass), though always slower than light, can be made to chase a photon arbitrarily closely. From this we shall deduce that the conformal and projective structures of space-time are *compatible*, in the sense that every $\mathcal{C}$ -null geodesic is also a $\mathcal{P}$ -geodesic. We call a manifold M endowed with a compatible pair of $\mathcal{C}$, $\mathcal{P}$ structures a *Weyl space*.

A Weyl space $(M, \mathcal{C}, \mathcal{P})$ possesses a unique *affine structure* $\mathcal{A}$ such that $\mathcal{A}$ -geodesics coincide with $\mathcal{P}$ -geodesics and $\mathcal{C}$ -nullity of vectors is preserved under $\mathcal{A}$ -parallel displacement. Conversely, the existence of such an $\mathcal{A}$ for a pair $(\mathcal{C}, \mathcal{P})$ implies that $\mathcal{C}$ and $\mathcal{P}$ determine a Weyl structure. In view of this theorem, one may say that light propagation and free fall define a Weyl structure on space-time; we symbolize the latter by $(M, \mathcal{C}, \mathcal{A})$.

(d) In a Weyl space, one can define an arc length (unique to within linear transformations) along any non-null curve (i.e. a curve whose tangent is nowhere null) k by requiring that the corresponding tangent vector $\mathbf{T}$ be congruent, at each point of k , to a non-null vector $\mathbf{V}$ which is parallelly displaced along k (congruence of vectors at a point is, of course, well defined in a conformal space [Pirani and Schild, 1966, p. 291]). Intuitively, this means that two infinitesimal line elements of k , ds_p and ds_q , situated at p and q respectively, are considered as congruent on k if the infinitesimal connection vector belonging to ds_q arises from that associated with ds_p by parallel transport from p and [sic: to] q , followed by a rotation (or pseudorotation) at q (pp. 65–67). ... Applying this to the time-like world line of a particle P (not necessarily freely falling), one obtains a proper time ($\equiv$ arc length) t on P , provided two events on P have been selected as zero point and unit point of time (pp. 67–68). ...

(e) It is now a straightforward matter to formulate additional assumptions that are necessary and sufficient in order that a Weyl space $(M, \mathcal{C}, \mathcal{A})$ be a Riemann space, in the sense that there exists a Riemannian metric $\mathcal{M}$ compatible with $\mathcal{C}$ (i.e., having the same null cones) and having $\mathcal{A}$ as its metric connection. The Riemannian metric is then necessarily unique, up to a constant positive factor ... we show that $(M, \mathcal{C}, \mathcal{A})$ is Riemannian if and only if the proper times t, t' on two arbitrary, infinitesimally close, freely falling particles P, P' are linearly related (to first order in the distance) by Einstein-simultaneity: i.e. if and only if whenever $p_1, p_2, \dots$ is an equidistant sequence of events on P (ticking of a clock) and $q_1, q_2, \dots$ is the sequence of events on P' that are Einstein-simultaneous with $p_1, p_2, \dots$ respectively, then $q_1, q_2, \dots$ is (approximately) an equidistant sequence on P' (p. 69).

We see that despite the stated divergences of the F & P & S method from Weyl's method, the F & P & S method is a *geodesic* method in essentially the same sense as Weyl's. Furthermore, Weyl as well as Ehlers *et al.* deny rigid rods any role in the physical foundations of the GTR ST-metric. Hence they all forfeit, no less than Synge does, such justification as the presumably *Pythagorean* infinitesimal metric behavior of rigid rods can furnish for adopting the Riemannian species of ST-metric.

Does the ontology of the F & P & S version of the geodesic method succeed in avoiding a counterpart in its logical edifice to the *deus ex machina* or 'manna' which its authors found objectionable in Synge's speciation of a Finslerian ST-metric into a uniquely Riemannian one? And have they supplied in *their* version of the geodesic method a non-arbitrary *terra firma* physical foundation for the GTR's adoption of the Riemannian kind of ST-metric, as contrasted with just adopting that kind of metric at the outset, as Weyl does in his version of the geodesic method? Let us deal with the latter question initially.

To answer this question, let us first recall what F & P & S wrote under items (d) and (e) of their summary of their reasoning. Under (d), they addressed themselves to what they call a 'Weyl space', i.e., to a manifold which is endowed with a compatible pair of conformal and projective structures. And they say that in a Weyl space one can define an arc length on any non-null curve k , to within a choice of unit and of origin, by specifying the conditions under which two (generally disjoint) infinitesimal line elements ds_p and ds_q on the non-null curve k 'are considered as congruent on k '. In particular, arc length can be defined here on k via "requiring that the corresponding tangent vector T be congruent, at each point of k , to a non-null vector V which is parallelly displaced along k " (1972, p. 67). Note that congruence is here relativized to a particular curve k , so that the congruence or incongruence of two given intervals will generally be *path-dependent*: Two intervals I_1 and I_2 each of which is part of two generally different curves C and C' can be congruent with respect to C while being incongruent with respect to C' . On the time-like world line of a (not necessarily freely falling) particle P , the arc length obtained after a choice of zero and unit is taken to be the 'proper time t on P '.

Now, under their (e), F & P & S state two (equivalent) conditions each of which is necessary and sufficient in order that a particular Weyl space

$(M, \mathcal{G}, \mathcal{A})$ be a Riemann space in the following sense: There exists a Riemannian metric $\mathcal{H}$, *unique modulo a factor k* , which has the same null cones as $\mathcal{G}$, and with respect to which the affine geodesics of $\mathcal{A}$ are likewise metrical geodesics. Specifically, E & P & S point out that this will be the case if and only if the *proper time* congruences which they defined on the given Weyl space $(M, \mathcal{G}, \mathcal{A})$ are *not* curve-dependent but are 'approximately' preserved under an Einstein-simultaneity mapping of the time-like world line of any freely falling particle P into the infinitesimally close world line of another arbitrary free particle P' . (Einstein-simultaneity is, of course, characterized by the value $\varepsilon = \frac{1}{2}$ in the Reichenbachian notation of chapters 12 and 20.)

At least one of the significant contributions made by the beautiful and illuminating E & P & S paper is to have demonstrated that compatible $\mathcal{G}$ and $\mathcal{P}$ structures permit the derivation of the existence of an affine connection $\mathcal{A}$ which is *unique* in the following sense: ' $\mathcal{A}$ -geodesics coincide with $\mathcal{P}$ -geodesics and $\mathcal{G}$ -nullity of vectors is preserved under $\mathcal{A}$ -parallel displacement' (E & P & S, 1972, p. 67). But their language in (d) may misleadingly generate *unwarranted* doubts in regard to the 'manna vs. *terra firma* issue' I have posed. I shall formulate these unfounded doubts presently by reference to two different Weyl spaces $W_1(M_1, \mathcal{G}_1, \mathcal{P}_1)$ and $W_2(M_2, \mathcal{G}_2, \mathcal{P}_2)$ whose respective *unique* affine connections are $\mathcal{A}_1$ and $\mathcal{A}_2$ and whose properties are the following: (1) The interval congruences which E & P & S defined on W_1 by their $\mathcal{A}_1$ -parallel transport prescription are *path-dependent*, and (2) the interval congruences they defined on W_2 by their $\mathcal{A}_2$ -parallel transport prescription are *path-independent*, so that W_2 is a Riemann-space rather than merely a Weyl space.

It *might* seem that however 'natural' their prescription for the introduction of interval congruences into a given Weyl space, E & P & S are no less 'manna-laden' than Weyl for the following reasons:

(1) The particular path-dependent congruences which E & P & S have thus introduced into W_1 are NOT *dictated* by the latter's compatible structures $\mathcal{G}_1$ and $\mathcal{P}_1$; nor – for that matter – is the introduction of any congruences at all into the space, and

(2) The compatible structures $\mathcal{G}_2$ and $\mathcal{P}_2$ possessed by W_2 *do not dictate* the particular path-independent interval congruences which the E & P & S transport prescription has imposed on W_2 , congruences which

do in fact select for W_2 a Riemannian metric g that is unique up to a positive constant factor k .

But John Stachel has explained why it would be an error to think that E & P & S employ as much 'manna' as Weyl's outright assumption of the Riemannian character of the ST-metric: In their (d), E & P & S innocuously employ nominal definitions to introduce an arc length which has been made available in all but name by the physical and logical resources of their prior construction, much as a quadrilateral whose opposite sides are parallel is an entity of Euclid's axiom system even *before* it is *called* 'parallelogram'. Hence E & P & S are entitled to regard Weyl's outright assumption that the GTR ST-metric will be Riemannian in kind as *less motivated physically* in the context of the conformal and projective structures of *his* geodesic method than it is in the E & P & S version of that method.

It would seem that in order to make good on their claim of having provided this stronger physical motivation, E & P & S need *NOT* regard their generation of a Riemann ST-metric via their prescription for introducing a *proper time* t into a Weyl space-time like W_2 as tantamount to *supplementing* the $\mathcal{G}_2$ and $\mathcal{P}_2$ structures by reliance on a *special class* of *physical clocks*! It is, of course, in principle a question of experimental fact whether, in any given case, the physical construction of E & P & S yields a Riemann space-time or only a Weyl space-time. Hence with regard to clocks that do read the proper time t which they have introduced into Weyl spaces, E & P & S (1972, p. 69) write: "Taking P, P' as world lines of (not necessarily freely falling) particles, one can easily see that the Riemannian property means that whenever two standard clocks (as determined above) associated with P, P' have equal rates at p , then they also have equal rates at q ." And they point out (*ibid.*) that this "can, in principle, be tested experimentally."

The otherwise admirable presentation by E & P & S expositively glosses over points that make it less perspicuous than it should be if its similarities and differences from Weyl's version of the geodesic method are to be readily discernible. Thus, in Weyl's 1921 presentation, it is stated axiomatically (1968 reprint, p. 195) that a (non-vanishing) Riemannian metric tensor g_{ab} can be *so chosen* that the ST-directions dx^i of *photons* at any given worldpoint e satisfy an equation of the form $g_{ab} dx^a dx^b = 0$. In the E & P & S presentation, no metric structure at

all is assumed at the outset. But for a metrically non-committal tensor g_{ab} , they *prove* as a Lemma (1972, pp. 72–73) that the tangent vector $\mathbf{T}$ at e of any light ray L through e satisfies an equation of the form

$$g_{ab}\mathbf{T}^a\mathbf{T}^b=0.$$

Conversely, any such vector $\mathbf{T}$ is a tangent vector of some light ray. By contrast to Weyl, who had *assumed* the existence of a Lorentzian conformal structure whose null directions are the directions of light rays, E & P & S have shown that a few empirically well-motivated properties of light propagation suffice to *construct* such a structure. Their optical derivation of a conformal structure can be likened to Helmholtz's derivation of a Riemannian metric by requiring the free mobility of rigid bodies. (The function g introduced in Axiom L_1 by E & P & S is merely a proof-technical tool.)⁶

Relying on this comparison of the relative merits of the Weylian and E & P & S versions of the geodesic method in regard to the provision of a physical motivation for the GTR ST-metric, we can now comment on whether the E & P & S construction has a counterpart to the *deus ex machina* which E & P & S found objectionable in Synge's method. Here I must conclude that I was unable to discern assumptions which have the physical status of 'manna' in their construction, as does Synge's speciation of a Finslerian metric into a Riemannian one.

I should add that I do not see that from either the ontological or epistemological point of view, Woodhouse (1973) has improved on the E & P & S construction, as he claims to have done. For example, in regard to primitive concepts, Woodhouse takes as fundamental a light signal between two events rather than a light ray in space-time, writing (1973, p. 496): "One important advantage of this approach over that of Ehlers, Pirani, and Schild is that no assumption is made, *ab initio*, about the paths in space-time along which light signals propagate: These paths are deduced from statements about the emission and absorption [*sic*] of light." Yet this purported improvement appears to be dependent not only on requiring given emission and absorption events to belong to one and the same light signal, but also on the time-asymmetrical distinction between emission and absorption. What then is the *net* gain over the E & P & S concept of a light ray between a pair of events? In the same

⁶ The role of g in L_1 is just to characterize the degree of smoothness of the light rays.

vein, contrast E & P & S's statement that they "do not introduce any intrinsic difference between future and past here" (1972, p. 72) but only their oppositeness or relative difference, with Woodhouse's avowedly axiomatic assumption "that on each particle there is a natural arrow of time" (1973, p. 495).

So much for our comparison of the E & P & S method with other constructions.

We must now devote some attention to the logical and epistemic status of the concept of a *free* massive particle in the geodesic method as such.

In his 1921 paper, Weyl appeals to Einstein's principle of equivalence of inertial and gravitational mass as the basis for the physical assumption of his geodesic method that a freely falling massive particle has a determinate inertial motion in the sense of a determinate time-like ST-trajectory, regardless of its other characteristics such as mass and (chemical) composition. Says he (Weyl, 1968 reprint, p. 196; translation is mine): "In the theory of relativity, the projective and conformal structure have an immediately palpable significance. The former, the inertial tendency of the ST-direction of a moving material particle which imparts to it a determinate 'natural' motion upon its release in a given ST-direction, is that unity of inertia and gravitation by which Einstein replaced both, but which has so far lacked a suggestive name." Since Weyl first introduced his geodesic method, it has become known that upon its release in a given ST-direction, the ST-trajectory of a freely falling gravitationally *multipole* particle will not be the same as that of a gravitationally *monopole* particle which is falling freely under otherwise relevantly identical conditions. And if the geodesic method is to yield the results demanded by the full-blown GTR, the time-like ST-trajectories of free massive particles having a gravitational *multipole* structure *cannot* count as time-like *geodesics*, whereas the world-lines of the monopole particles should so count.

But John Stachel has pointed out that this complication threatens the construction of the ST-metric by the geodesic method with logical circularity, and derivatively with epistemological circularity, as we shall now see. Let us assume, for argument's sake, that one has succeeded in assuring that the massive particles are free to the extent that they are

For a more detailed statement of this point by Weyl, see his (1923b, p. 219).

effectively insulated from any and all KNOWN influences which, according to the GTR, would alter their ST-trajectory and render the geodesic method inoperative. This is explicitly assumed to be feasible by Graves (1971, p. 171). And it is recognized as a practical problem by Kundt and Hoffman (1962, p. 305), who speak of the emission of 'geodesic test particles' by an observer and remark: "This will pose practical problems since the effects of spin motion and of the electromagnetic field must be made negligible." Furthermore, let us assume that even if there are as yet unknown perturbing influences in our GTR universe, our massive particles are effectively free of them as well by some good fortune. Even then the *multipolarity* of spinning particles poses more than practical problems!

Specifically, the geodesic method is faced by the following problem: It must be able to avail itself, in a logically *non*-circular way, of the distinction between gravitationally monopole and multipole free massive test particles. For it must *exclude* the ST-trajectories of the latter when stipulating that only the world lines of the former are to be coextensive with the time-like geodesics of the desired Riemannian metric. And, of course, the latter metric is first going to be made available logically in the geodesic method by combining this geodesicity stipulation with the conformal structure of light rays. Hence no part of the GTR which is predicated on the resulting metric may be presupposed, on pain of vicious logical circularity, in order to distinguish at the outset between the two species of free test particles when imposing the geodesicity requirement on only one of them. Yet, as Stachel has explained, it would appear that precisely such a presupposition is needed, much as for method (i) of p. 731.

This logical circularity seems to beset the E & P & S version of the geodesic method no less than Weyl's. Consequent upon this logical circularity, there is also the epistemological one of knowing how to identify without (tacit) appeal to the resulting metrical theory, the gravitationally *right kind* of free particle, when first trying to ascertain the metric by the geodesic method. The geodesic method would become vitiated by patent epistemic circularity, if it were to seek to identify gravitationally monopole free particles by first attempting to *ascertain* whether their ST-trajectories are, in fact, geodesics! That Weyl was quite generally sensitive to the risk of possible epistemic circularity is attested by the fact that he wrote as follows (1923a, p. 19): "Thus, if in the real world

it is possible for us to discern the propagation of effects, and of light propagation in particular, and if moreover we are able to recognize as such the motion of free mass points which obey the metrical field, and to observe that motion, then we are able to read off the metric field from this alone, without reliance on clocks and rigid rods" (translation is mine).

As we noted, Graves (1971, pp. 170–171) was likewise aware of the risk of epistemological circularity in the geodesic method. But he overlooked the problem of logical *and* epistemic circularity posed by gravitational multipolarity of free particles. And he took no cognizance of the role of the conformal structure in the geodesic method but assumed without argument that this method can achieve its objective by means of the projective structure alone, writing (Graves, 1972, p. 648): "The choice of geometry, and the geodesic hypothesis, is derived from the central empirical conclusion from the equivalence principle: a class of space-time paths is observationally singled out as the trajectories of *all* freely falling bodies. The geometry is selected so that these are its geodesics, not because of some procedure with special rods and clocks." Having thus greatly oversimplified the ontological and epistemic merits of the geodesic method vis-à-vis both the method of rods-and-clocks and Synge's chronometric method, Graves felt entitled to write (1971):

... Reichenbach, Grünbaum, and their disciples are correct in arguing against the conventionalists that once a standard has been chosen, the geometry is determined [reference omitted]. However, they completely miss the point by speaking as if the standard were some *particular* kind of rods and clocks. The crucial fact is rather that *any* body will do. Furthermore, it is not any assumed metrical properties of the standard (such as invariance under transport) that are relevant, but simply its path in free fall. Once we can trace all these paths we have all the information we need to determine the geometry. We have not yet *measured* the intervals and curvatures, to be sure; but measurement is a separate operation. In measurement we *discover* the metrical features of the already determined geometry; whereas on Reichenbach's theory, we *invent* a geometry as an attempt to make the results of our various measurements consistent (p. 172).

But, in view of multipolarity, it is *not* the case, as Graves claims, that 'any [free] body will do'. And as against Graves' oversimplified verdict as to the relative merits of the geodesic method (1971, pp. 152–164 and 170–172), it seems that our analysis sustains a different conclusion: Each of the three methods (i), (ii) and (iii) above has its own logical and epistemological 'dirty linen', as it were. On p. 731, I stated the likely moral of this fact!

In this latter vein, John Stachel has commented on the E & P & S version of the geodesic method, writing (1973):

Since this method of test particles has recently come into favor, it is very satisfying to see it fully developed. Of course, each of the three methods currently proposed for explicating the significance of the metric structure of space-time (rods and clocks, paths of massive particles and clocks, and paths of massless and massive particles) has advantages and drawbacks. Thus, it is perhaps better to regard them as alternative ways of looking at the implications of a metrical structure for space-time than to claim absolute superiority for one (p. 292).

Coupled with Graves' neglect of the geodesic method's specified 'dirty philosophical linen' is his blithe assumption that ontologically *this* method must be preeminently foundational not only for standard 1915 GTR, but especially for the Riemannian metric of Wheeler's *empty* curved space-time. To assess this assumption, I shall now disregard the aforementioned 'dirty linen' when I ask: In an empty GMD universe, what is the ontic status of the massive free test particles of the geodesic method? Clearly, in the GMD ontology, these geodesic material particles are *not* primitive entities, any more than are the much maligned rods-and-clocks of method (i) or than Weyl's photons or Synge's atomic clocks. Instead, in the empty world of GMD, the geodesic material particles are themselves held to be literally constituted out of empty, curved, metric space-time to begin with, no less than any and all other traditionally used physical metric devices! It is a consequence of GMD's all-out *geometrical* reductionism and absolutism that the metric of empty space-time can no more *first* be *induced* in the ST-manifold (in part or whole) by geodesic particles than by rigid rods. In other words, the GMD ST-metric cannot consistently first be grounded ontologically even partly on the inertial behavior of geodesic particles, although that behavior *can* physically *realize* or single out ST-trajectories which qualify as (time-like) geodesics with respect to the GMD ST-metric. For in GMD no entities other than the structure of empty space-time itself are *ontologically* necessary for endowing space-time with metric ratios or with such curvature properties as are determined by these ratios. Hence, according to GMD, any and all particles or radiation that serve to specify the Riemannian metric of the empty GMD space-time can do so at best *only epistemically* as a means of our discovering that metric.

Surely it cannot legitimately be assumed tacitly without any further ado that in the context of GMD the highly derived or non-primitive

ontological status of the geodesic particles can be rendered innocuous or mitigated in this context by simply appealing to their smallness qua so-called 'test' particles. Why then should we accept Graves' undaunted declaration that vis-à-vis both the rods-and-clocks of method (i) and the atomic clocks of method (ii), the freely falling material geodesic particles of Weyl's method (iii) play a preeminently foundational role *ontologically*, not only in standard 1915 GTR but in GMD??

So far, we have only partially articulated the ontic status of the metric in GMD as distinct from the ontologically more non-committal standard 1915 GTR. We must now turn to the ramifications of that articulation both for the metric itself and for the curvature properties of empty space-time.

3. THE ONTOLOGY OF EMPTY CURVED METRIC SPACE IN THE GMD OF CLIFFORD AND WHEELER

The empty space(time) of the Clifford-Wheeler GMD is both metric and curved. We shall deal with these two fundamental entities in turn.

(a) *The Ontology of the METRIC of Empty Curved Space-Time in GMD*

I have explained why I cannot share Graves' somewhat dogmatic belief that the geodesic method enjoys foundational ontic preeminence both in standard 1915 GTR and in GMD. I shall nonetheless find it very useful to employ Weyl's geodesic method as a framework for treating the issues which will now concern us in this Section (a). Wheeler's GMD assumes a Riemannian kind of ST-metric at the outset no less than Weyl's version of the geodesic method does. Hence the latter is fully as germane to our ontological analysis of GMD as the E & P & S version of that method. Furthermore, I believe that *mutatis mutandis*, the conclusions which we shall reach by reference to Weyl's geodesic method are obtainable not only via the E & P & S version of the geodesic method but also via either the rods-and-clocks method or Synge's chronometric method. These forthcoming ontological conclusions fully allow but do not require that there is *concordance* of the metric results furnished by the three major metrical methods (i), (ii) and (iii).

It is, of course, a matter of experimental fact whether there is such concordance, i.e., whether these three methods are actually alternative

specifications and or physical realizations of one and the same ST-metric modulo k . Thus, Kundt and Hoffman (1962) have written:

In the general theory of relativity, the metrical tensor can be pathwise determined by making measurements with rigid rods and standard clocks [reference omitted]. However, rigid rods and atomic clocks are devices obeying the laws of electrodynamics and quantum theory. The fact that the metrical tensor has a gravitational-inertial as well as a metrical significance means that standard length and standard time are also determined by the inertial motions of free particles of both non-zero and zero rest masses. To be more precise: the projective structure of space-time given by the (timelike) non-null geodesics, together with the conformal structure given by the null lines determine the metrical structure to within a scale factor that is independent of position [reference omitted].

The question therefore arises whether gravitational standard time and non-gravitational standard time are the same, a question which can only be answered by experiment (p. 303).

Accordingly Kundt and Hoffman devoted their paper to suggesting “an experiment by means of which, at least in principle, gravitational time can be measured within any degree of accuracy (allowed by the uncertainty principle)” (p. 303).

We shall see that two quite different verdicts will be reached as to the ontic status of the GMD ST-metric according as one rejects or accepts one of the cardinal tenets of Riemann’s own conception of the foundations of the metric of continuous n -dimensional physical space. Indeed, the latter tenet will be seen to be incompatible with the ontological commitments of GMD. In view of the important ramifications of this incompatibility, it behooves us now to recall this explicit and fundamental idea of Riemann’s, which pertains to the basis for the metric equality or ‘congruence’ of intervals in a continuous n -dimensional spatial or temporal manifold. Weyl (1922, pp. 97–98) discussed this idea and then expressed it very concisely as follows (1922, p. 101): “... according to Riemann, the conception ‘congruence’ leads to no metrical system at all, not even to the general metrical system of Riemann, which is governed by a quadratic differential form.” As the reader will recognize from pp. 495–498 of Ch. 16, here Weyl is alluding to the fact that in Riemann’s 1854 Inaugural Dissertation, an important distinction is drawn between two kinds of metrics as follows: A (non-trivial) metric which is ‘implicit’ in, or intrinsic to, the space on which it is defined, on the one hand, and a metric which is correspondingly non-implicit or extrinsic. As was explained in Ch. 16 (p. 501), this distinction of Riemann’s emphatically must NOT be confused with the very different *Gaussian* contrast that is

unfortunately also expressed by means of the terms 'intrinsic' and 'extrinsic'. Since Riemann set forth his distinction only illustratively and intuitively, I have attempted to give a relatively much more precise explication of it in Ch. 16 (Part A, §2), a chapter which is reprinted from Grünbaum (1970).⁸

On the basis of this explication, I gave a more precise formulation (p. 527) of the following thesis of Riemann's, which I dubbed 'Riemann's Metrical Hypothesis' (RMH), (pp. 498–501) and whose attribution to Riemann I documented there in detail. In a continuous n -dimensional physical space, there is no kind of *intrinsic* basis at all for any non-trivial metric that would qualify as *implicit* in that space to within a constant positive scale factor k . Hereafter I shall refer to this latter hypothesis by the abbreviation 'RMH'.

Let us pause briefly to illustrate from two authors the intellectual ubiquity of the intuition underlying Riemann's claim that the presumed differentiable manifold of physical space is devoid of any non-trivial 'implicit' or intrinsic metric. Thus, writing in the opening paragraphs of his book on Einstein's theory, Max von Laue (1952) declared without referring to Riemann:

A continuum does not carry a metric within itself – whereas in the case of a chain, one can arrive at a metric possessed by it by numbering its links, nothing like that obtains in the case of an undifferentiated thread, so that – if we wish to metrize the latter, we must put a measuring device with its subdivisional markings upon it (p. 1, translation is mine).

And then von Laue (1952) adds:

Likewise the three-dimensional continuum, which we call space is – initially devoid of any metric partitioning, [and] possesses no geometry of its own (p. 2, translation is mine).

In the same vein, the mathematician Morris Kline spoke of the metric as *imposed* on the continuous spatial manifold when he wrote (1972, p. 892): "Strictly speaking, Riemann's curvature, like Gauss's, is a property of the metric imposed on the manifold rather than of the manifold itself." Note here the contrast between "a property ... of the manifold itself," i.e., a property intrinsic to the manifold, on the one hand, and, on the other hand, "a property of the metric imposed on the manifold," i.e., a property which is extrinsic to the manifold because it is first bestowed upon it by an imposed and hence extrinsic metric.

⁸ For a summary of that paper going beyond the Abstract in Ch. 16, see Grünbaum (1971a).

It is a corollary of RMH, as extended to continuous 4-dimensional physical space-times, that entities *extrinsic* to these differentiable manifolds are needed ontologically to generate, in a GTR universe of given mass-energy distribution, the GTR's *particular* metric ratios of ST-intervals, ratios by which the metric geometry is determined to be what it is in the given case. By contrast, according to GMD, the space-time in which we live is both empty *and* metric (Riemannian), so that *no* entities other than 4-dimensional empty space-time itself are *ontologically* necessary for endowing that physical manifold with a particular set of Riemannian metric equalities! Thus, in virtue of the emptiness of the ST-manifold, its Riemannian metric must be 'implicit' in it or intrinsic to it modulo k in the sense of at least *not* being imposed on it, but of being grounded solely in the very structure of that 4-dimensional physical manifold itself.

Furthermore, as we saw at the end of §2, according to GMD's all-out *geometrical* reductionism and absolutism, it cannot be held that the ST-metric is *FIRST induced* in its 4-dimensional manifold by the behavior of such entities as (atomic) clocks, rods, light rays, geodesic particles or the like. For though in GMD these agencies are claimed *not* to be 'foreign entities immersed' in space-time – to use Wheeler's (1962a, p. 361) parlance – they *are* asserted to be *reducible* to the ontologically and logically *prior* metric geometry of space-time. Yet *Riemann* deemed the behavior of at least some devices of this kind to be *ontologically necessary* for first conferring a metric structure upon continuous physical space or time. Since contrariwise, GMD holds all such devices to be themselves ontologically reducible to curved empty metric space-time, in that theory the ST-metric cannot be ontologically grounded on their behavior. Instead, in GMD such devices can at best *physically realize* the intrinsic metric equalities with which GMD claims empty space-time to be endowed. In this way, such devices can function *epistemologically* in measurement.

We see that if GMD's program of reducing all of physics to metric ST-geometry were to be successful empirically – a possibility which Wheeler himself has now discounted as unlikely – then this putative explanatory success could redound to the empirical *discredit* of RMH. Conversely, in the absence of such massive empirical success or pending such success, the assumption of RMH may be warranted and has the

consequence of impugning the ontology of GMD. In my (1970) and in Ch. 16 (pp. 501–504), I pointed out this logical consequence after having followed Weyl (1922, §12) in calling attention to the prior scientific fruitfulness of RMH for standard 1915 GTR.

Thus in the Abstract of Ch. 16 (p. 450) and in Grünbaum (1970, p. 470), I stated that I question the Clifford-Wheeler statements of GMD in regard to “the compatibility of the theory [GMD] with the Riemannian metrical philosophy apparently espoused by its proponents.” The exegetical aspects of this expression of incompatibility will receive detailed attention within subsection (b) later on below. But I trust that the logical incompatibility between the GMD ontology of the metric and RMH is now sufficiently clear. Hence I do not see that my earlier claim of such incompatibility deserves the following assessment made by Earman (1972): “Some philosophers are not content with mere empirical objections and can be satisfied only by a-prioristic refutations. In this vein, Adolf Grünbaum has argued that GMD is incoherent” (p. 646). According to Earman, my argument for incompatibility is ‘a prime example’ (1972, p. 647, fn. 13) of what he had described as “the tendency (all too prevalent in current philosophy) to view relativity theory as only a source for grist for some philosophical mill” (1972, p. 634). And Earman deplored the latter pernicious tendency at the end of the very same sentence which he began by extolling Graves’s (1971) book in the following way: “I will not dwell here on ... how loudly I applaud his [Graves’s] attempt to describe relativity theory from the point of view of scientific realism” (Earman, 1972, p. 634). Presumably, on Earman’s view the espousal of scientific realism is free from the fetters of philosophical apriorism because a scientific realist interpretation of the GTR is vouchsafed in a straightforward way, much as one can read off the names from a telephone directory. I shall not enter here into objections to such uncritical ontological *literalism* for physical theory. Suffice it to remind the reader of how much analysis it took in Ch. 19 to try to show that the coarse-grained entropy of classical statistical mechanics may be construed in scientific realist fashion instead of being a mere anthropomorphism.

Just as the presumed truth of RMH can serve to impugn the GMD ontology of the metric, so also it can provide a basis for dealing with the comment made by Glymour (1972, p. 338) on my question (Ch. 16, p. 467, and Grünbaum, 1970, p. 487) as to “what serves to individuate

the metrically homogeneous punctal event elements of the space-time manifold?" in GMD. If the aforementioned fruitfulness of RMH for classical 1915 GTR is taken to warrant the adoption of RMH as a working assumption, then it provides precisely the grounds for which Glymour asked when he said: "If metric relations are specifically excluded from the class of individuating relations then ... we must ask for the grounds of this exclusion" (1972, p. 338). For, according to RMH, no intrinsic non-trivial metric relations are available, and surely the individuality of the fundamental world-point entities of the manifold cannot be made to rest on *extrinsic* metrical relations.

Glymour asked for the grounds for excluding metric relations from the class of individuating relations after having made the following statement by reference to a certain version of Leibniz's principle of the identity of indiscernibles: "we may then regard the entities of the manifold as individuated by their metric relations" (*ibid.*). To this I say the following. If the empirical success of GMD's program of reduction were sufficient to sustain its dictum '*Physics is geometry*' (Wheeler, 1962b, p. 225), then RMH could no longer serve as a viable basis for impugning the GMD ontology of the metric. And *if*, moreover, the intrinsic metric relations postulated by GMD could be demonstrated to individuate the world-points of its ST-manifold, then I would agree with Glymour that the doubts I raised about GMD in regard to individuation are unwarranted. Indeed, Glymour overlooked that in my (1970), I had said as much (see Ch. 16, p. 504) à propos of the so-called 'intrinsic coordinate systems' which obtain under conditions sufficiently heterogeneous to yield four *scalar* fields that can individuate *each* world point.

In any case, if such philosophical appraisal of GMD as inquiring into the adequacy of its principle of individuation is to be condemned with Earman as either invidiously aprioristic or 'frivolous' (Earman, 1972, p. 647), then one wonders what role, if any, he envisions for philosophical analysis and criticism of a scientific theory, as distinct from purely empirico-mathematical or technical appraisal of it.

Returning to the upshot of our discussion of the GMD ontology of the metric so far, we can say the following:

(a) If we accept the GMD claim that modulo k the 4-dimensional ST-manifold is *intrinsically* and Riemannianly metric, then we must reject Riemann's own RMH.

(b) GMD postulates that there is an intrinsic basis for the metrical nullity of photon lines in space-time *and* for a particular set of metrical ratios among non-null ST-intervals. GMD then asserts that free massive particle trajectories qualify – via $\oint ds = 0$ – as time-like geodesics with respect to any metric ds that generates both these intrinsic metric ratios and the stated metrical nullity for photons. Thus, in GMD the metrical time-like geodesicity of massive particle world-lines and the metrical nullity of light rays obtains as a matter of intrinsic fact. Having this presumed intrinsic foundation, the *geometrical* status of these two sets of world lines in GMD is *not* a matter of convention, stipulation or human decree in the sense of these terms discussed in Ch. 16, §3.

(c) Since GMD asserts massive free particles and photons to be *geometrically reducible* entities, the geodesic method of Weyl and of E & P & S does not provide an ontological foundation of the GMD ST-metric but only a specification or circumscription of that metric modulo k .

Being mindful of this item (c), let us recall from our discussion of the rods-and-clocks method under (i) in §2 above Einstein's reservations about the foundational invocation of rigid rods as a (partial) physical basis for the GTR ST-metric. Then it becomes clear from item (c) that *qua ontological foundation* of the ST-metric, the geodesic method should probably be deemed even less satisfactory *for GMD* than rigid rods (and clocks) are for classical 1915 GTR.

I have stressed the logical incompatibility of the GMD ontology of the ST-metric with RMH. But I have yet to articulate the ramifications for the appraisal of that GMD ontology, if one assumes RMH along with founding the ST-metric of classical (standard) 1915 GTR on the geodesic method. As a corollary of my impending statement of these ramifications, I shall be able to comment further on Graves's (1971) ontological account of the GMD ST-metric vis-à-vis both Hans Reichenbach's ST-philosophy and some of my own views.

Assuming RMH and the philosophical import of RMH developed in Ch. 16, §3, our earlier scientific statement of Weyl's geodesic method in §2 (iii) of this chapter now requires philosophical supplementation in the form of a series of statements as follows:

(1) Given the assumption of RMH, there is *no* intrinsic basis and hence no *initial* GEOMETRICAL *warrant* for singling out any particular classes of ST-trajectories as being respectively distinguished in re-

gard to time-like *metrical geodesicity* and *metrical nullity*. Nor is such an initial geometrical warrant created by the mere pre-geometrical fact that the physical behavior of photons and of relativistically free massive particles – *pace* the gravitationally multipolar ones – does respectively single out two classes U_p and U_m of ST-trajectories and thereby determines the membership of their union U . For, given RMH, the objective physical determinateness of the membership of U does not detract one iota from the fact that humans selected U without intrinsic geometrical warrant, as against other classes of ST-trajectories *on the foundation of which a differently curved or even flat ST-structure would arise*.

(2) The physically determinate membership of U assures the compatibility of the following stipulations laid down by the geodesic method: An indefinite Riemann metric (tensor) *is to be* SO CHOSEN that with respect to *one and the same such metric* ds

(a) The photon trajectories at any given world point are to BECOME *metrically null*.

(b) The ST-trajectories belonging to the proper subclass U_m of U are to be TURNED INTO time-like GEODESICS via the intra-theoretic defining equation $\delta \int ds = 0$.

Graves, who confines his attention to the subclass U_m in this context, distinguishes with commendable clarity between the physical determinateness of the membership of U_m , on the one hand, and the geometrical status of that membership as time-like geodesics on the other, when he says (1971, p. 171): “These paths are real and observable, quite independent of any geometry in which they may be represented.” Indeed, the crucial logical transition from the mere physical ST-trajectorihood of the members of U_m , which is pre-geometric, to their geometrical status becomes evident when Graves says of the free massive particle trajectories (1972, p. 648): “The geometry [i.e., the ST-metric modulo k] is selected so that these are its geodesics.”

Given the assumption of RMH, the members of U_m are metrical time-like geodesics by human convention in the sense of Ch. 16, §3: the metrical ratios of non-null intervals with respect to which the members of U_m so qualify are devoid – in Riemann’s sense of RMH – of any ‘implicit’ or intrinsic foundation. Similarly for the metrical nullity of the photon trajectories which constitute the subclass U_p of U . Thus, if RMH is true, all the physical trajectories belonging to U acquire their specific *geo-*

metrical status in Weyl's method extrinsically by human convention.

(3) It follows that if RMH is true, then the metrical structure of space-time — and thereby the very constitution of the only autonomous substance in the GMD monistic ontology — depends crucially for being what it is not only on the physically-determinate membership of U , but also on the *intrinsically*-UNFOUNDED, humanly-stipulated ascriptions of metrical *geodesicity* and *nullity* to the appropriate members of U ! As the reader may recall from Ch. 16, §3 (especially p. 550) and as is argued in greater detail in Grünbaum (1968, Ch. III, §§3 and 6) and in Massey (1971), the inevitable metrical conventionality consequent upon RMH is not at all of the merely trivial semantical kind which holds alike for any and all as yet semantically uncommitted vocabulary, including the as yet unpreempted WORDS or *noises* 'geodesic', 'metric ds', etc. For the latter trivial semantical kind of conventionality obtains alike, regardless of whether the ascriptions of metrical geodesicity and nullity made by Weyl's geodesic method do have an intrinsic foundation, as GMD maintains, or fail to have such a foundation, as RMH claims. By contrast, if the GTR ST-metric does have the intrinsic kind of foundation claimed by the GMD ontology, then the stated non-trivial metrical conventionality does *not* inevitably obtain.

GMD asserts that all of physics, even if not psychology, is reducible to the metric geometry of space-time. I submit that whatever one's views on the mind-body problem, attributes depending fundamentally, even if only partly, on particular human stipulations are not at all the kind of item which can reasonably be essential ontological elements in the very constitution of the only autonomous substance in the physical world. Yet as we saw, if RMH is true, precisely this is the case in the monistic geometrical GMD ontology: Granted RMH, human stipulations enter ontologically — not just verbally! — into making the metric geometry of space-time be what it is, and thereby these stipulations paradoxically generate the character of the only autonomous 'physical' substance recognized by GMD. For on the assumption of RMH, human conventions are indispensable to the GMD metric geometry in the sense of entering essentially into generating those metrical properties of ST-intervals by which the GMD geometry is made to be uniquely what it is modulo k in our actual world.

This brings me to some further remarks on Graves's account of the

GMD ontology of the ST-metric. Graves (1971, pp. 170–171; 1972, p. 648) refers to the statement that the ST-trajectories belonging to U_m are (time-like) geodesics as ‘the geodesic hypothesis’. The ontological credentials of this ‘hypothesis’ are quite different, however, according as RMH is true *or* the geodesicity asserted by the hypothesis does have an intrinsic foundation, as claimed by GMD. For in the former case, the geometrical status of the members of U_m as time-like geodesics has no intrinsic foundation but is humanly stipulated. One would therefore have expected that qua exponent of the GMD ontology, Graves would unequivocally opt for *denying* that geodesicity is parasitic on human stipulation in the context of GMD’s geometrical reductionism and absolutism.

More positively, one would have expected Graves to disavow the wide-ranging metric permissiveness countenanced by the conventionalists in favor of saying something like the following: (1) According to GMD, the metrical time-like geodesicity of the members of U_m (and the metrical nullity of the members of U_p) has an intrinsic foundation, and (2) the possession of this intrinsic foundation distinguishes the metric ST-geometry specified by Weyl’s geodesic method from an infinitude of others which lack such a foundation but are equally countenanced philosophically by the conventionalist adherents of RMH. Instead, we find Graves writing as follows (1971):

Therefore, what the geodesic hypothesis does is to select out of the infinitely many possible geometric manifolds that particular one whose geodesics coincide exactly with the world lines of freely falling test particles, and to proclaim this as the real geometry of the space-time region.

Now the conventionalists are right in saying that we are in no sense *required* to make this choice of geometry (p. 171).

Presumably GMD postulates that in our actual ST-world there is *only one intrinsic* metric geometry modulo k and that Weyl’s geodesic method specifies that distinguished metric geometry G_i . If this postulate be granted, then it is at least misleading to say with Graves that what the geodesic method does is ‘to proclaim’ G_i ‘as the real geometry of the space-time region’. For supposedly the latter ascription of ‘reality’ to G_i should codify not merely the logically *pre*-geometric physical determinateness of the membership of U_m (and of U_p); it should also render the ontological thesis of GMD that our actual world’s ultimate substance

is a space-time endowed with G_t as its *only intrinsic metric geometry*. And in that case, the unique status of G_t 'as the real geometry' of space-time is hardly a matter of 'proclamation'. Furthermore, in that case both Graves and the conventionalists are then alike mistaken in claiming, as Graves (1971, p. 171) would have it, that "we are in no sense *required* to make this choice [G_t] of geometry." For, if the GMD ontology of the ST-metric is true, there surely is a sense in which we are required to make this choice of geometry, since G_t is distinguished by having an intrinsic foundation. And presumably we would wish our geometric theory to render this unique ontological anchorage. Graves seems to gloss over the difference between the pre-geometric meaning of 'real' and the geometric meaning here, which is synonymous with 'intrinsically-based' or 'implicit' in Riemann's sense.

We have already given reasons at the end of §2 for questioning Graves's belief that vis-à-vis the devices of the *other* methods, the geodesic particles of Weyl's geodesic method play a preeminently foundational role in the ontology of the ST-metric not only in classical 1915 GTR but in GMD. We shall now see, furthermore, that even if the geodesic method did enjoy such foundational preeminence, the following criticism by Graves (1971) of Reichenbach and me — in which Graves uses 'GR' as an abbreviation for 'General Relativity' — is unjustified:

It is vitally important to recognize that this identification of space with matter and geometry with physics is the central conceptual feature of GR. Such philosophers as Reichenbach and Grünbaum have misinterpreted the theory by attempting to keep them separate. They have been concerned with such questions as "Once we have laid down a set of 'co-ordinative definitions' (even if by arbitrary convention), to what degree is the geometry uniquely determined, and once this is done, what is the determinate form of the physical laws in this geometry?" While these questions are interesting in their own right, the whole approach is antithetical to that of GR (p. 152). Neither the conventional character of the choice of coordinative definitions *nor* the possibility of using a particular stipulation to make an empirical determination of the geometry is really relevant for GR, though it may be important for an understanding of space in general (apart from any particular theory in which the term 'space' appears). For a given physical configuration, GR does single out a unique geometry, but this uniqueness results from the identification of physical with geometrical entities, *not* from a choice of coordinative definitions (p. 153).

Graves's two groups of remarks here invite several replies. It will be useful to begin with his second set of claims.

(1) In his paper 'Graves on the Philosophy of Physics', Howard Stein (1972) wrote:

The part of Graves's conceptual analysis of the technical structure of general relativity that I have found most disappointing is his discussion of the problem of measurement. He is rather severely critical of what may be called the received account of geometrical and chronometrical measurement, that associated with the names of Reichenbach and Grünbaum; but the grounds of his criticism, and the content of his own view, seem to me not very clearly expressed. For instance, he tells us (153) that for a given physical situation "GR does single out a unique geometry, but this ... results from the identification of physical with geometrical entities, *not* from a choice of coordinative definitions." But "identification of physical with geometrical entities" would seem to be just what Reichenbach *meant* by 'coordinative definition' (p. 631).

(2) Weyl's geodesic method poses and answers precisely the following kind of question which Graves deemed alien to GTR when he objected to Reichenbach's having asked: Once we have laid down a set of 'coordinative definitions' (even if by arbitrary convention), to what degree is the geometry uniquely determined? For Weyl showed that once we have laid down the coordinative definition that light rays are to count as metrically null lines of the theory, the physical ST-metric, and thereby the metric geometry, is determined only up to conformal transformations of the metric tensor. By the same token, Weyl showed that the coordinative definition which correlates the time-like geodesics of the theory with free massive particle trajectories is itself also insufficient to determine the physical metric modulo k , since that correlation yields only the projective structure without the conformal one vouchsafed by the metrical nullity of photons. Thus, Weyl also answered the question as to how far the geometry is uniquely determined by the mere imposition of the coordinative definition that the ST-trajectories in U_m are to count as time-like geodesics.

Graves might have appreciated this point, if he had not overlooked the role of the conformal structure in the geodesic method, when he assumed without argument that this method can specify the ST-metric modulo k by means of the projective structure alone. Indeed, as we saw at the beginning of §2 (iii), mathematical theorems pertaining to the positive definite metrics of ordinary surfaces and 3-spaces tell us that *geodesic* mappings of such spaces onto themselves do *not* determine their positive definite Riemann metrics modulo k . And more generally, theorems concerning geodesic mappings (Laugwitz, 1965, §13; Eisenhart, 1949, §§40 and 41) lend themselves at once to answering the question as to the extent of the determination of the metric geometry of a *physical* 3-space or space-time, *if* a certain family of *physical* paths is held to qualify geometri-

cally as geodesics. Mathematical physicists, including relativists, can and do ask, as well as answer, this kind of question without necessarily being beholden to Reichenbach's theory of coordinative definitions. How then can concern with such questions on the part of Reichenbach and myself be a type of inquiry whose 'whole approach is antithetical to that of GR', as Graves (1971, p. 152) alleges it to be? It is regrettable in this connection that when Graves (1971, p. 153n) cites Putnam's (1963) as support for his views, he takes no cognizance of the arguments in my detailed 'Reply to Putnam' (1968, Ch. III, and 1969, pp. 1-150).

For had Graves taken such cognizance (e.g., of Grünbaum, 1968, pp. 272 and 276), some of his misunderstandings of Reichenbach might have been obviated. Thus we saw that Graves had complained that the Reichenbachian account of geometrical conventions is not really relevant for GTR, "though it may be important for an understanding of space in general (apart from any particular theory in which the term 'space' appears)" (Graves, 1972, p. 153). But in so complaining, Graves overlooks the following documented point, which I made in my 'Reply to Putnam' (1968, p. 276): "Despite emphatic statements by both Reichenbach and me to the contrary ..., Putnam assumes that when we offer what we consider to be a stipulational specification of congruence ..., we do so utterly unmindful of the physical validity of such principles as Riemann's concordance assumption [for rigid transported rods and standard clocks] i.e., blind to the theoretical context of the physics in which the congruence specification is to function."

In the context of Weyl's geodesic method, it is therefore fully germane to GTR to ask the Reichenbachian kind of question in regard to the extent of the determination of the geometry by specified coordinative definitions. Moreover, once the metric geometry is uniquely determined, it is likewise germane to GTR to go on and ask, as Reichenbach and I have done, "what is the determinate form of the physical laws in this geometry?" For if we use Weyl's geodesic method *foundationally* for the metric geometry as countenanced by Graves, then a partial answer to the *latter* question is given by adducing none other than Einstein's field equations of the GTR! These equations assert a lawful connection between the metric tensor, as specified by Weyl's method, and the matter-energy distribution.

(3) Graves (1972, unpublished) has replied to Howard Stein's aforementioned defense of Reichenbach (Stein, 1972, p. 631) as follows:

... I still insist on the main point of disagreement with Reichenbach. According to the equivalence principle, *all* freely falling bodies, regardless of their composition and any possible use as rods or clocks, will follow the same observable space-time trajectories in a pure gravitational field. And in any Riemannian manifold a class of paths can be singled out mathematically as its geodesics. The 'geodesic hypothesis' simply puts together these empirical and mathematical facts: the geometry whose geodesics coincide with the paths of freely falling bodies is taken to be the 'true' geometry. This identification is derived from an empirical fact, and not from criteria or conventions chosen a priori.

Did Graves portray Reichenbach's views accurately here when he depicted Reichenbach as holding that the identification effected by the geodesic hypothesis "is derived ... from criteria or conventions chosen a priori"? To see that the answer is negative, we need only cite Reichenbach's account of the status of the metric of 3-space as grounded physically on rigid rods. Reichenbach writes (1958):

This analysis reveals how definitions and empirical statements are interconnected. As explained above, it is an observational fact, formulated in an empirical statement, that two measuring rods which are shown to be equal in length by local comparison made at a certain space point will be found equal in length by local comparison at every other space point, whether they have been transported along the same or different paths. When we add to this empirical fact the definition that the rods shall be called equal in length when they are at *different places*, we do not make an inference from the observed fact; the addition constitutes an independent convention. There is, however, a certain relation between the two. The physical fact makes the convention unique, i.e., independent of the path of transportation. The statement about the uniqueness of the convention is therefore empirically verifiable and not a matter of choice (pp. 16-17).

... *It is again a matter of fact that our world admits of a simple definition of congruence because of the factual relations holding for the behavior of rigid rods; but this fact does not deprive the simple definition of its definitional character* (p. 17; italics in original).

In the light of Reichenbach's statement, we can ask Graves in what sense the identification of the physically determinate members of U_m as time-like ST-geodesics of the geometric theory can be claimed with Graves to be 'derived from an empirical fact'? Is it not clear from our analysis that Graves's claim of *derivability* of the so-called geodesic hypothesis from Einstein's equivalence principle is unsound? And is it not also clear that by characterizing the assertion made by the geodesic hypothesis as a convention, Reichenbach does *not* claim this convention to have been 'chosen a priori'? Here Graves overlooked the point of pp. 122-123.

In the published part of his 'Reply to Stein and Earman', Graves had also alleged the *derivability* of the geometry from the equivalence principle via the geodesic hypothesis, when he declared (1972, p. 648): "The choice

of geometry, and the geodesic hypothesis, is derived from the central empirical conclusion from the equivalence principle: a class of space-time paths is observationally singled out as the trajectories of *all* freely falling bodies. The geometry is selected so that these are its geodesics, not because of some procedure with special rods and clocks." Yet Graves himself seems to have implicitly *denied* the *derivability* of the geometry from the equivalence principle when he wrote (1971, p. 171): "Now the conventionalists are right in saying that we are in no sense *required* to make this choice of [GMD] geometry." For, according to him, the equivalence principle permits the 'derivation' of the geodesic hypothesis, *and* again, according to Graves — once the geodesic hypothesis is accepted, the geometry is determined. Thus on Graves's account, once the equivalence principle is accepted, the geometry is no longer open to choice. But Graves does accept the equivalence principle while claiming as well that "we are in no sense *required* to make this [GMD] choice of geometry."

As Reichenbach saw it, presumed empirical facts about the behavior of rigid rods assure their *consistent* interchangeability as congruence standards. But these facts do *not* thereby assure the derivability of the congruence specification furnished interchangeably by initially coinciding transported rods. Analogously, as explained by Weyl and further by E & P & S, presumed empirical facts about the behavior of massive and massless particles assure the *compatibility* of the projective and conformal structures as requirements governing one and the same Riemannian metric or affine connection. But, as I see it, these presumed empirical facts do not assure *in and of themselves*, as Graves explicitly believes, the derivability of the metric (geometry) modulo *k*.

(b) *The Ontology of the CURVATURE of Empty Space in GMD*

As was mentioned in subsection (a) of the present §3, in my reply (Ch. 16, p. 450, and Grünbaum, 1970, Abstract, p. 470) to "A Panel Discussion of Grünbaum's *Philosophy of Science*" (*Philosophy of Science* 36 [Dec. 1969]), I had questioned the Clifford-Wheeler statements of GMD in regard to "the compatibility of the theory [GMD] with the Riemannian metrical philosophy apparently espoused by its proponents." Thereafter, Clark Glymour (1972) devoted his paper 'Physics by Convention' to a critique of Chapter 16 and Grünbaum (1970). Speaking of the latter essay, Glymour (1972) writes in the opening paragraph of his paper:

Grünbaum's replies to his critics, especially his most recent reply (1970) ... involve unusually important claims which fail to be buttressed by the arguments he gives. I have in mind such claims as ... that the foremost advocates of geometrodynamics, Clifford and Wheeler, were and are enmeshed in contradiction. My own view is that all of these claims are dubious or false, but I shall be less concerned with establishing their falsity than with discrediting the arguments offered for them (p. 322).

Glymour (1972) amplifies this assessment near the end of his paper as follows:

Grünbaum charges the advocates of geometrodynamics, Clifford and Wheeler in particular, with inconsistency. Since these men have maintained that matter reduces to curved space, they must also have thought, according to Grünbaum, that curvature is an intrinsic property of space. Yet, Grünbaum argues, both Clifford and Wheeler deny that space has an intrinsic metric. But this is inconsistent, Grünbaum concludes, since "... this curvature would need to obtain with respect to a metric implicit in empty space" (Grünbaum, 1970). Now I do not think this claim especially important, partly because I am not at all convinced that Wheeler would deny that space has intrinsic metric properties, and partly because the program of geometrodynamics certainly does not require such a denial. Even so, I doubt that Grünbaum has provided, or can provide, anything like sufficient grounds for his conclusion. He gives no argument at all as to why we should think curvature properties presuppose or require or 'would need to obtain with respect to' a metric. It cannot be because the curvature tensor of space does in fact determine a *unique* Riemannian metric, for that is not true, as Grünbaum himself appears to have noted [at this point, Glymour cites Grünbaum (1963, pp. 89–105)]. Perhaps by 'curvature' Grünbaum intends properties some of which are not determined by the curvature tensor alone: affine properties generally, perhaps, or sectional curvatures. But a 3-dimensional manifold fitted with a Riemannian connection does not, in general, have a *unique* compatible metric, even up to similarity (p. 338, italics added). Even if the properties in question include all affine properties and sectional curvatures it is not clear that they determine a *unique* metric (p. 339, the *second* italics are added). ... So interpreted, then, Grünbaum's contention that curvature requires metric is at best moot. Of course, Grünbaum may simply have meant that curvature properties are just not the sort of thing that can exist unsupported. But he has given us no shade of reason why that might be so, let alone demonstrated that what is required for their support is a metric (p. 339).

Here Glymour raises two issues as follows:

(i) Quite *apart* from how Clifford and Wheeler conceived specifically of curved empty space as the building material of the physical world, do curvature properties of space presuppose (require) a metric, so that it would be inconsistent to claim that space is intrinsically curved but is devoid of any intrinsic metric?

(ii) What is the answer to the latter question of presupposition and inconsistency when posed in the specific context of the assumptions made by Clifford and Wheeler, and if the answer is positive, did either Clifford or Wheeler avow such an inconsistency?

(i) *Question (i): Do curvature properties of empty space require a metric?* Do curvature properties of space require or presuppose a metric, so that it would be inconsistent to claim that space is both intrinsically curved and yet devoid of an intrinsic metric? We shall see before long that the correctness of an affirmative answer to this question turns crucially on whether the given curvature properties include those specified (or determined) by the *covariant* 4th rank Riemann-Christoffel curvature tensor or are *confined* to those furnished (determined) merely by the *mixed* Riemann-Christoffel curvature tensor (which is contravariant of rank 1 and covariant of rank 3). But before I can develop this point, I must show in some detail how Glymour misconstrued our question (i) from the outset.

Glymour overlooked that, to begin with, the issue here is (α) whether given curvature properties require or depend on SOME Riemannian metric OR OTHER, and *not*, as Glymour would have it, (β) whether given curvature properties require or presuppose a metric which is *unique* at least to within a choice of the unit (i.e., *unique* up to a constant positive factor k or 'similarity'). Glymour's conflation of (α) with (β) is important here for the following two reasons: 1. If the answer to (α) were *affirmative*, that answer to (α) would be logically *weaker* than an affirmative answer to (β), since the former merely asserts that *some* Riemannian metric *or other* is presupposed, while the latter asserts that a metric *unique* at least up to a constant positive factor k is presupposed, and 2. if the answer to (α) were affirmative, that *weaker* affirmation would be *sufficient* to justify the charge of *inconsistency* against the claim that space is both intrinsically curved and yet devoid of any intrinsic metric at all.

Let me refer to the claim that space is both intrinsically curved and yet devoid of *any* intrinsic metric as 'MUC' to convey its assertion of the feasibility of the Metrical Unsupportedness of intrinsic Curvature. And let me call the charge that MUC is *inconsistent* 'The Inconsistency Charge'. The strong affirmation that the answer to (β) is positive asserts that given curvature properties require a metric *unique* at least up to the factor k , and hence this strong assertion will hereafter be designated as 'the curvature requirement of a unique metric' or as 'CRUM'. In terms of these designations, I maintain that the two reasons I gave for rejecting Glymour's conflation of (α) with (β) show the following: It was unavailing for Glymour to have tried to undermine the inconsistency charge by wrongly assuming that this charge against MUC rests on the logically

strong affirmative answer CRUM to question (β) rather than merely on the logically weaker affirmative answer to (α). As we noted above, the latter weaker affirmation asserts merely that given curvature properties require *some* Riemannian metric *or other*. On the basis of his erroneous belief that the inconsistency charge against MUC rests on the strong assertion CRUM of uniqueness, Glymour then proceeded to argue irrelevantly by means of certain technical results that this uniqueness claim is either false or gratuitous, *even if* the given curvature properties comprise some which are *not* determined by a 4th rank curvature tensor alone.

It is instructive to note the reason which prompted Glymour to be concerned to assert the falsity or unfoundedness of CRUM with respect to a set of given curvature properties that is *wider* than one containing only those determined by the 4th rank curvature tensor alone. In the first edition of this book (see Ch. 3, Section B above), I had given several examples to show that even in two-dimensional Euclidean space – where the one independent component of the curvature tensor vanishes along with the total Gaussian curvature K – the vanishing curvature K of that 2-space *fails* to determine a Riemannian metric or metric tensor which is *unique* up to a choice of unit. Thus, I had *denied* the uniqueness claim CRUM *with respect to the Riemannian (Gaussian) curvature of 2-space*.

In personal correspondence, I had called Glymour's attention to this denial of mine in response to the following statement by him in an earlier draft of his (1972) which he had kindly made available to me: "Even assuming he is accurate in the views he ascribes to Clifford and Wheeler, Grünbaum would *only* be correct [in his charge that MUC is inconsistent] if a curvature tensor on a differentiable manifold somehow determined or required a *unique* metric tensor; but that is not true" (*italics added*). As shown by my quotation above from his *published* text, Glymour modified his cited earlier statement to the extent of taking cognizance of the fact that I had denied CRUM with respect to the curvature tensor. But he persisted in his earlier misguided concern with discrediting CRUM, albeit by now widening the membership of the class of curvature properties to which this uniqueness claim is held to pertain. Instead of realizing that the inconsistency charge against MUC need not rest on any kind of uniqueness claim CRUM, and that I had never claimed it does, Glymour mistakenly remained convinced that it must. Having thus adhered to his initial misconstrual (β) of our question (i), he therefore proceeded to

inquire irrelevantly though interestingly what wider miscellaneous assortment of curvature properties might perhaps sustain the uniqueness claim CRUM. Not surprisingly, even then he found CRUM wanting or at best moot. But we saw that this finding does not at all undermine my inconsistency charge against MUC. So much for Glymour's own handling of our question (i).

We must now deal with this latter question when properly construed as initially posing the issue (z) whether given curvature properties require (presuppose, depend on) *some* Riemannian metric *or other*. And let us propose to construe the given curvature properties as those specified or determined by the 4th rank Riemann-Christoffel curvature tensor. We must then point out at once that in the context of our present inquiry the following fact is very important indeed: There is an ambiguity in the expression '*the* Riemann-Christoffel curvature tensor' as used nowadays, according as we mean the so-called Riemann tensor of the *first* kind – which is the *covariant* 4th rank curvature tensor – or the so-called Riemann curvature tensor of the *second* kind, which is the *mixed* curvature tensor that is contravariant of rank 1 and covariant of rank 3.⁹ Clearly, this ambiguity makes it dependent on the context whether given statements about curvature properties, construed as rendered by '*the* Riemann tensor', do refer to or involve a commitment to both kinds of Riemann tensor or to only one of them!

A few brief preliminaries will provide us with the means to define both kinds of Riemann tensor. Let us denote ordinary partial differentiation by a comma followed by the variable with respect to which we differentiate partially. Thus, the partial derivative $\partial g_{ik} / \partial x^m$ of the symmetric covariant, second-rank metric tensor g_{ik} would be written as ' $g_{ik,m}$ '. The component g 's of this metric or 'fundamental' tensor are functions of the coordinates subject to the restriction that their determinant be non-zero. In this notation, the so-called Christoffel symbols of the first kind are defined by

$$[ij, k] = \frac{1}{2}(g_{ik,j} + g_{jk,i} - g_{ij,k}).$$

⁹ For a statement of these designations, see, for example, Eisenhart (1949, ch. I, §8). Speaking of the first and second kinds of Riemann curvature tensor, Morris Kline (1972) writes: "Either form is now called the Riemann-Christoffel curvature tensor" (p. 1127).

Furthermore, the contravariant conjugate metric tensor g^{lk} is defined by the condition $g_{ml}g^{lk} = \delta_m^k$, where we sum over the repeated index l , and δ_m^k is the Kronecker delta: $\delta_m^k = 1$ for $k = m$, and $\delta_m^k = 0$ for $k \neq m$. The so-called Christoffel symbols of the second kind are then defined by

$$\left\{ \begin{matrix} l \\ i \ j \end{matrix} \right\} = g^{lk} [ij, k].$$

Now consider a set of functions Γ_{ij}^l of the coordinates, such that these functions transform like the Christoffel symbols of the 2nd kind under a change of coordinate system but do *not necessarily* share all of the other properties of these Christoffel symbols.

A set of such functions Γ_{ij}^l is called a set of 'coefficients of affine connection'. *There are* affine connections Γ_{ij}^l which have a greater number of independent components than is possessed by any of the second kind

of Christoffel symbols $\left\{ \begin{matrix} l \\ i \ j \end{matrix} \right\}$, although all affine coefficients transform just like the latter Christoffel symbols. Specifically, in a space of n dimensions, there are affine connections which have as many as n^3 independent

components. But any Christoffel symbol $\left\{ \begin{matrix} l \\ i \ j \end{matrix} \right\}$ has *fewer* independent components, since the number of independent components possessed by any such symbol in an n -space is restricted by the merely at most $(n^2 + n)/2$ independent components which the two *symmetric* fundamental conjugate tensors g_{ij} and g^{lk} have both collectively and severally. Thus, whenever an affine connection Γ_{ij}^l has a greater number of independent

components than any Christoffel symbol $\left\{ \begin{matrix} l \\ i \ j \end{matrix} \right\}$, the partial differential

equations $\left\{ \begin{matrix} l \\ i \ j \end{matrix} \right\} = \Gamma_{ij}^l$ fail to have solutions g_{ij} and g^{lk} . Thus, *there are*

affine connections (i.e., sets of affine coefficients) which are *not* obtainable (derivable), via the set of Christoffel symbols of the second kind, from the combination of a fundamental metrical tensor g_{ij} with its conjugate.

Hence Christoffel's $\left\{ \begin{matrix} l \\ i \ j \end{matrix} \right\}$ are only a particular species of affine coefficients.

Although the affine connection is not itself a tensor, the following

object defined by means of it is a demonstrably mixed 4th rank tensor, contravariant of rank 1 and covariant of rank 3

$$(1) \quad R_{bcd}^a \equiv \Gamma_{bd,c}^a - \Gamma_{bc,d}^a + \Gamma_{n,c}^a \Gamma_{bd}^n - \Gamma_{n,d}^a \Gamma_{bc}^n,$$

where n is a dummy summation index and where we have adopted one of the two sign conventions used by various authors. Furthermore, we can lower the contravariant index in Equation (1) by contracting a metric tensor g_{an} with the particular mixed tensor R_{bcd}^a which corresponds to g_{an} in the following sense: R_{bcd}^n is generated via Equation (1) from those functions Γ_{an}^l that do qualify as Christoffel symbols $\left\{ \begin{smallmatrix} l \\ a \quad n \end{smallmatrix} \right\}$ with respect to the metric tensor g_{an} . Lowering the index in this way, we obtain

$$(2) \quad R_{abcd} = g_{an} R_{bcd}^n.$$

The *covariant* fourth rank tensor on the left-hand side of Equation (2) is the so-called Riemann (or Riemann-Christoffel) curvature tensor of the *first* kind, while the mixed 4th rank tensor on the left-hand side of Equation (1) is the so-called Riemann curvature tensor of the second kind.

In order to show now that the first kind of Riemann curvature tensor does presuppose a Riemannian metric while the second kind does not, let us note several prior relevant points.

(1) As we saw, there are affine connections Γ which are not derivable from the metric tensor via the Christoffel symbols of the second kind. Furthermore, the affine connections which are so derivable do *not* depend on the metric tensor for their definition (Adler *et al.*, 1965, ch. 2, esp. p. 48). Hence we shall say that Γ does not require (presuppose) a metric. Let us hereafter refer to any Γ_{ij}^l which is symmetric in its *subscripts* as 'symmetric' for brevity.

(2) And consider any tensor which is *constructible* or *obtainable* from any non-symmetric or symmetric affine connection Γ_{bc}^a in the following sense: The components of the tensor are expressible in terms of the Γ_{bc}^a and of their derivatives up to a certain order. Any such tensor is said to be a *differential concomitant* of the connection Γ_{bc}^a .

There is a known proof (Schouten, 1954, pp. 164–165) that the *covariant* Riemann curvature tensor R_{abcd} specified by our Equation (2) is *not* a differential concomitant of any non-symmetric or symmetric affine connection Γ alone; instead, R_{abcd} is a differential concomitant of the com-

bination of Γ with the metric tensor g_{ik} . Thus, we are entitled to say that the covariant Riemann curvature tensor is *not* a curvature tensor of the affine connection, although – as shown by Equation (1) above – the *mixed* Riemann tensor *is* such a curvature tensor. The stated non-obtainability of the covariant curvature tensor from Γ is not vouchsafed by the mere presence of the metric tensor g_{an} on the right-hand side of the equation $R_{abcd} = g_{an} R^a_{bcd}$, because that mere presence does not, of itself, prove the otherwise known fact that R_{abcd} fails to be *any* kind of differential concomitant of an affine connection.

Hence we can conclude that, at least with respect to currently known resources for constructing R_{abcd} , the covariant curvature tensor is simply *not defined without a metric tensor*. We shall see later in this subsection that this conclusion will be strengthened decisively by showing that, at any given point of space, (1) the components R_{abcd} of the covariant curvature tensor have the physical dimension L^2 of the square of *length*, where length will be seen to be a *metrical property*; and (2) the components R^a_{bcd} of the purely affine mixed curvature tensor have the dimensionless status of pure numbers, as we would expect from this *non-metrical* object. In this clear sense, the first kind of Riemann tensor requires or presupposes a Riemannian metric.

On the other hand, since the affine connection does not presuppose a metric, the mixed Riemann tensor, which is a differential concomitant of Γ alone in our Equation (1), does not require a Riemannian metric. Indeed it can be shown that the class of mixed Riemann curvature tensors which are obtainable in our Equation (1) from a metric tensor via the Christoffel symbols of the second kind is only a *proper* subclass of the set of mixed curvature tensors which are obtainable from any and all non-symmetric or symmetric connections Γ . R^a_{bcd} is a rather complicated differential concomitant of Γ in Equation (1). Hence the existence of mixed curvature tensors which are *not* obtainable from a metric tensor is not obvious from the mere fact that the Christoffel symbols $\left\{ \begin{smallmatrix} l \\ i \ j \end{smallmatrix} \right\}$, which are generated from the metric tensor, are only a particular *species* of affine coefficients. But the mixed curvature tensor “has certain symmetry properties only on the condition that the components of the affine connection are Christoffel symbols” of the second kind, and it then turns out that a mixed curvature tensor which is derivable from a metric tensor has fewer independent

components than one which is derivable from I 's that do *not* qualify as Christoffel symbols $\left\{ \begin{smallmatrix} I \\ i \ j \end{smallmatrix} \right\}$ (Bergmann, 1946, p. 169). Thus, the latter kind of mixed curvature tensor, which has the larger number of independent components, is not derivable from a metric tensor.

Our stated conclusions concerning the status of R_{abcd} and R^a_{bcd} respectively in regard to the presupposition of a metric are borne out by the following concise statement by Adler *et al.* (1965).

Note that, although we have introduced R^a_{bcd} for a Riemannian space, it is evident that our derivation holds also in a general affine space, since it involves only the coefficients of connection, and not the metric tensor itself. In the more general case, one need only replace $\left\{ \begin{smallmatrix} a \\ i \ c \ b \end{smallmatrix} \right\}$ by Γ^a_{bc} . However, as soon as we lower an index [by contraction with the covariant metric tensor] to form the tensor R_{abcd} , we commit ourselves to a metric space (ch. 5, p. 136; italics added).¹⁹

Note that the concluding sentence of this citation says that the covariant curvature tensor requires a metric space, i.e., *some* metric or other, but *not* that the curvature properties rendered by R_{abcd} presuppose a *unique* metric. It will be recalled that Glymour conflated these two different kinds of presupposition when he erroneously saddled me with needing the stronger of the two in order to establish my stated inconsistency charge.

As Morris Kline (1972, pp. 894–895, and 1125) points out in his recent monumental historical work, Riemann himself used the *covariant* tensor R_{abcd} to render curvature properties in his 1861 *Pariserarbeit*, which elaborates on his foundational 1854 Inaugural Dissertation. And Kline stresses that curvature tensor's presupposition of a metric by writing: "Strictly speaking, Riemann's curvature, like Gauss's, is a property of the metric imposed on the manifold rather than of the manifold itself" (Kline, 1972, p. 892), a lucid declaration which we had occasion to cite in §3 (a) as grist to my mill in more ways than one. In making this statement, Kline is well aware that the much later 1917–1918 work of Levi-Civita, Hessenberg, and Weyl yielded a purely affine generalization of Riemannian metric space such that the *mixed* curvature tensor can be construed non-

¹⁹ For notational convenience, I have replaced the Greek indices used by these authors by our Latin ones, and it will be noticed from the minus sign in their statement that they use rather unusual conventions such that $\left\{ \begin{smallmatrix} a \\ b \ c \end{smallmatrix} \right\} = -\Gamma^a_{bc}$; see also their pp. 48–50, esp. p. 50, Equation (2.10).

metrically (Kline, 1972, pp. 1131–1132 and 1134). The statement that the covariant curvature tensor involves a commitment to a metric holds even for that generalization of this tensor which corresponds to the *non-symmetric* metric tensor g_{ik} employed in Einstein's 'Relativistic Theory of the Non-Symmetric Field' (Einstein, 1955).

We can now formulate the two-fold upshot of our analysis for the answer which we shall give to our question (i).

In the first place, a given *covariant* curvature tensor field R_{abcd} over a particular manifold M *does require* some Riemannian metric on M , although one and the same such curvature properties can be conferred on M by metrics *differing* other than by a choice of unit. Thus, the manifold M on which the given R_{abcd} is defined does *not merely happen* to be a metric space. Nor can the given R_{abcd} on M be said to require some metric or other *just because* that curvature tensor qualifies as a differential concomitant of one or more metric tensors g_{ik} , respectively corresponding to as many Riemannian metrics any one of which may have been imposed on M . Instead, the curvature properties on M rendered by the given R_{abcd} presuppose a Riemannian metric because their very definition involves some metric tensor g_{ik} or other, so that these curvature properties would not be defined without having one of the metrics whose respective fundamental tensors g_{ik} can each generate the given curvature properties.

In the second place, a given *mixed* curvature tensor field R^a_{bcd} over a particular manifold M does *not* presuppose (require) a Riemannian metric, because R^a_{bcd} is *definable* on M without any metric tensor g_{ik} even when that curvature tensor does qualify as a differential concomitant of one or more tensors g_{ik} . Thus, even when a given R^a_{bcd} over a manifold M is a differential concomitant of the metric tensor of a metric whose imposition on M has turned M into a Riemann space S , we can say the following: The particular curvature properties R^a_{bcd} no more require that M be a metric space than the n -dimensionality of M presupposes the metric whose imposition has turned M into an n -dimensional Riemann space S .

These results now permit us to answer our question (i). If the curvature properties of a manifold M *include* those rendered or determined by the first (covariant) kind of Riemann curvature tensor R_{abcd} and are not *confined* to those specified by the second (mixed) kind of Riemann tensor, then the curvature properties of space M do require (presuppose) some

Riemannian metric or other. And in that case, it would indeed be *inconsistent* to claim that space is both intrinsically curved and yet devoid of any intrinsic metric. On the other hand, if the curvature properties of a manifold M are *solely* those specified or determined by a *mixed* curvature tensor, then no metric is presupposed and hence it is then obviously *not* inconsistent to assert MUC of M by saying that M is intrinsically curved but devoid of any intrinsic metric. In that case, the inconsistency charge against MUC would be plainly unsound. Indeed, if M 's curvature properties *exclude* the 'covariant' ones R_{abcd} while including the 'mixed' ones R^a_{bcd} , then M does not even qualify as a Riemann metric space but is merely an affinely connected space without a metric. The reason is that if M were a Riemann space rather than just an affine space, then its metric would be *sufficient* to assure, via the metric tensor, that M is endowed with 'covariant' curvature properties R_{abcd} after all no less than with the 'mixed' ones R^a_{bcd} . Conversely, if M is merely a space of affine connection without a metric rather than a Riemann space, then M 's curvature properties do exclude the covariant ones while including the mixed ones.

Thus, in any Riemann space, the metric of that space is both necessary and sufficient for *covariant* curvature properties R_{abcd} though only sufficient for mixed ones R^a_{bcd} . It follows importantly that in any Riemann space, the *metric* of that space is *non-trivially required specifically* by the *covariant curvature properties* with which that space is *automatically* endowed, rather than being just trivially required by the *metric character* of Riemann space!

We had seen earlier that since curvature properties are being construed as rendered by 'the' Riemann curvature tensor, the ambiguity of that construal as between R_{abcd} and R^a_{bcd} makes it inevitably *context-dependent* whether given statements about curvature pertain to a set of properties which properly includes the class R_{abcd} or pertain only to the restricted class R^a_{bcd} . And we see now that this context-dependence of what is intended by the term 'curvature' is tantamount to whether the manifold whose curvature properties are being discussed is a Riemann space or only an affinely connected space.

Apart from his conflation of question (α) with question (β) above, Glymour's critique of my inconsistency charge against MUC is vitiated by his neglect of the following facts: (1) The *context* in which my inconsistency charge against MUC occurred is indispensable for determining

whether the curvature properties to which I referred in that charge include the covariant ones R_{abcd} or not. (2) On the very page [Chapter 16, p. 503 and Grünbaum (1970, p. 523)] on which I asserted – as the basis for my inconsistency charge – that curvature properties require a metric, I explicitly referred to a ‘Riemannian manifold’ and to ‘the framework of Riemannian geometry’ and *not* to a merely affine space! – as the context to which my assertion of metrical presupposition pertained, and (3) In the avowedly Riemannian metrical context of my inconsistency charge, the covariant curvature properties R_{abcd} , which do presuppose a metric, are indeed included among the curvature properties to which that charge pertains.

It follows that, contrary to Glymour, the inconsistency charge against MUC as leveled by me is true.

Furthermore, it is evident that qua intended criticism of the views I set forth in Chapter 16 and Grünbaum (1970), the following statement by Earman (1972, p. 647, n. 13) is directed against a straw man: “Clark Glymour ... has argued against the claim that curvature properties must obtain with respect to a metric.” I had never denied that *there exists* a species of curvature properties which do not require a metric, namely the mixed ones R^a_{bcd} , since I had *not* claimed that *any and all* curvature properties must obtain with respect to a metric. Nor was it essential in the context of my argument – which avowedly pertained to ‘the framework of Riemannian geometry’ – to point out that the particular curvature species R^a_{bcd} need not have a metrical anchorage. By contrast, as shown by Glymour’s discussion of the consistency of MUC, he construed curvature properties as those rendered by ‘the’ Riemann curvature tensor, *but* took no cognizance at all of the *relevant* fact that the covariant curvature properties R_{abcd} must obtain with respect to a metric.

Finally, we can appraise the following statement of Glymour’s (1972):

Of course, Grünbaum may simply have meant that curvature properties are just not the sort of thing that can exist unsupported. But he has given us no shade of reason why that might be so, let alone demonstrated that what is required for their support is a metric (p. 339).

As is patent from our analysis, it is indeed the case, contrary to Glymour, that “curvature properties are just not the sort of thing that can exist unsupported.” Any given covariant curvature properties R_{abcd} over a manifold M in fact cannot exist as such unsupported by some metric or other,

though a multiplicity of Riemannian metrics are each capable of conferring these given properties R_{abcd} on M . Moreover, any given mixed curvature properties R^a_{bcd} over a manifold M cannot exist as such unsupported by some affine connection Γ or other, though a multiplicity of affine connections are each capable of endowing M with these given properties R^a_{bcd} . Glymour fails to see that there is *abundant* reason – rather than ‘no shade of reason’ – for holding that curvature properties as such cannot exist unsupported, because he is again victimized by his conflation of two different requirements as follows: On the one hand, requiring the support of a *unique* metric for R_{abcd} and of a *unique* Γ for R^a_{bcd} , and, on the other hand, requiring more weakly the support of respectively *some* metric or other, and of *some* Γ or other.

In a paper entitled ‘Space-Time Indeterminacies and Space-Time Structure’, presented to the Boston Colloquium for the Philosophy of Science in 1973, Glymour has offered a rebuttal to my arguments.

Glymour considers the following triplet of contentions (in footnote 8 of this paper of his):

- (1) Space-time has, intrinsically, certain curvature properties
- (2) Space-time does not have, intrinsically, a metric
- (3) Material bodies reduce to curved space-time

Glymour then goes on to reply to me as follows (*ibid.*, fn. 8):

Suppose someone were to offer a space-time theory strictly in accord with 1–3, and suppose he intended to include among curvature properties some which are not obtainable from an affine connection: the chief examples of the latter are sectional curvatures. His theory, then, would perhaps contain variables for an affine connection, various scalar vector or tensor fields, and quantities, such as the [mixed] (1, 3) curvature tensor definable from these. We can also imagine that his theory contains a field quantity which takes every pair of vectors in every tangent space at every point into a real number in a smooth way, and that he calls this quantity the ‘sectional’ curvature. We can even imagine that his theory contains a quantity which is a (0, 4) tensor [i.e., a 4th rank *covariant* tensor], and that his axioms guarantee that this quantity has all of the symmetry properties of a (0, 4) curvature tensor, and in addition all of the properties of R_{ijkl} (e.g., action on vector fields) that can be stated without use of the metric tensor. His theory does not contain any explicit quantity which is a metric tensor.

Of course, no one has actually developed such a theory, but someone could I suppose, and I think it would be entirely reasonable to regard the properties he is talking about as what we ordinarily regard as curvature properties. When is it reasonable to say that someone holding such a theory is committed to the existence of a metric? My inclination is to say that he is so committed when his axioms are strong enough to permit the definition of a metric quantity from the curvature and affine quantities to which he is already committed. That was the point of my observation, against Grünbaum, that even all affine

properties and sectional curvatures do not contain a unique metric, and so do not permit the definition of a metric.

One can object against such a theory that it is incomplete, that there are relations among the different curvature properties that cannot be described without a metric, and that might (or, conceivably, might not) be true, but I fail to see how such a theory, coupled with claims 1–3, is inconsistent.

This criticism gives me the opportunity to offer what I believe to be decisive further support for my claims by now showing that, at any given point, the components of the *covariant* curvature tensor (0, 4) have the physical dimension L^2 of the square of length – where length is a METRICAL property! – while those of the *mixed* (1, 3) curvature tensor have the dimensionless status of pure numbers. It is also going to be relevant that, as is well known, the reciprocal L^{-2} of L^2 is the physical dimension of the Gaussian curvature K at a point of a surface, and that L^{-2} is also the physical dimension of the sectional curvature K_N at a point with respect to a given orientation N . I shall now proceed to demonstrate the stated results for the two curvature tensors in order to show that they render Glymour's purported speculative counterexample to me untenable.

Let me introduce the general considerations which are to follow by a simple example. On the surface of a 2-sphere embedded in Euclidean 3-space, let the two sets of surface coordinates be respectively the colatitude ϕ in *radians* and the longitude θ in *radians*. By the very definition of radians, the values of ϕ and θ are each dimensionless pure numbers, as are the coordinate differentials $d\phi$ and $d\theta$. If a is the length of the radius of the sphere, then the familiar spatial metric on the sphere is given by

$$ds^2 = a^2 d\phi^2 + a^2 \sin^2 \phi d\theta^2.$$

Since the physical dimension of ds^2 is L^2 , each of the two terms in the sum on the right-hand side must have the dimension L^2 . But since $d\phi^2$ and $d\theta^2$ are each pure numbers, it is clear that the metric coefficients g_{11} and g_{22} must each be of dimension L^2 , as indeed they are because a^2 is the square of a length. In particular, in the case of the *unit* sphere, where a^2 and hence g_{11} has the value *unity*, g_{11} is of dimension L^2 . Thus, at least in our example, the metric coefficients g_{ik} are not only *qualitatively* of dimension L^2 but contain a metric scale factor whose value is dictated by the choice of the *unit* of length! More generally, it is clear that after any particular coordinatization is introduced in an n -dimensional manifold, alternative metrizations which differ *only* in the choice of unit of length

will issue in metric tensors g_{ik} which, in the *given* coordinate system, differ only in the scale factor contained by them. Thus, the choice of metric unit of length always makes itself felt in the functions g_{ik} !

As shown by our simple example of the coordinate system (ϕ, θ) , *there are* permissible coordinate systems in which we *MUST* assign the dimensional status L^2 to the metric coefficients g_{ik} ! Moreover, since a coordinate system is specified by a function from points into ordered n -tuples of pure real numbers in any space of n -dimensions, it is always *possible* to assign the dimensional status L^2 to the metric coefficients of *any* coordinate system, even in the simple case when one or more of them has the constant value unity.

Thus, to take another simple example if (ρ, θ) are polar coordinates on the Euclidean plane, not only the values of θ in radians but also those of ρ can qualify as pure (dimensionless) real numbers. In that construal, the values of $d\rho^2$, ρ^2 and $d\theta^2$ are each pure numbers. And we can then make perspicuous the dimensional role of unit coefficients of dimension L^2 in the components g_{ik} by writing the metric in the form

$$ds^2 = 1 \cdot d\rho^2 + 1 \cdot \rho^2 d\theta^2.$$

We see that it is possible to impose the requirement that the dimensional status of the metric tensor g_{ik} be *the same* in *every* coordinate system not only qualitatively but also in the quantitative sense of containing a scale factor which is dictated by the choice of unit. At least for reasons of simplicity and generality, I shall impose this requirement in the general statement which I shall give below. But we shall note at the conclusion of our analysis of the two curvature tensors that it suffices for the refutation of Glymour's purported counterexample that **THERE ARE** coordinate systems in which at least some of the metric coefficients not only **MUST** be of dimension L^2 but also **MUST** contain the metrical scale factor dictated by the choice of unit of length! Since the general formulations which are about to follow are to be applicable to 4-dimensional space-time, ds will need to refer to space-time 'length', and ' L ' will denote the latter physical dimension. For our purposes, I shall not need to take account of the separate dimensions of spatial length and temporal duration in order to accommodate the cases in which space-time can be split up into space and time. But it is to be understood that I allow for this refinement in the general statement which I shall now proceed to give.

A coordinate system is specified by a function from points into ordered n -tuples of real numbers. The real number coordinates clearly have the dimensionless status of *pure* numbers. Hence, the coordinate differentials dx^i and dx^k , and products of these, are also pure numbers. But in the equation for the metric ds specified by $ds^2 = g_{ik} dx^i dx^k$, ds^2 has the dimension L^2 . Hence at any given point of space, the components of the metric tensor g_{ik} have the dimension L^2 while the components of its conjugate, the contravariant metric tensor g^{kl} , have the dimension L^{-2} (Schouten, 1951, p. 129). Partial derivatives of g_{ik} with respect to the coordinates have the same dimensions as g_{ik} , and similarly for g^{kl} .

It follows from the definition of the Christoffel symbols of the second kind given earlier in this section that, at any given point of the space, the entities represented by them are each dimensionless *pure* numbers.

For note that the $\left\{ \begin{smallmatrix} l \\ i \ j \end{smallmatrix} \right\}$ are obtained by contracting the contravariant metric tensor g^{lk} , whose components each have dimension L^{-2} , with the Christoffel symbols $[ij, k]$ of the first kind, which are objects of dimension L^2 , being one half of sums of partial derivatives of components of the tensor g_{ik} .

Since the dimensionless Christoffel symbols of the second kind are a particular species of affine coefficients, at any given point the coefficients Γ_{ij}^l are pure, dimensionless numbers, regardless of whether they satisfy a set of equations $\left\{ \begin{smallmatrix} l \\ i \ j \end{smallmatrix} \right\} = \Gamma_{ij}^l$ by being obtainable from the two conjugate metric tensors g^{lk} and g_{ik} .

These results will now enable us to see the following via our earlier Equations (1) and (2) of this section: Whereas at any given point, the components of the *mixed* curvature tensor (1, 3) are dimensionless pure numbers, the components of the covariant curvature tensor (0, 4) each have the dimension L^2 , where length is a *metrical property*. No wonder, therefore, that a purely affine space has the ontological resources to constitute the mixed curvature properties R_{bcd}^a out of *bona fide* non-metrical entities. But it is then likewise clear that, contrary to Glymour's allegation, the speculative (0, 4) tensor of Glymour's gleam-in-the-eye future non-metrical theory cannot possibly render the L^2 -dimensional properties R_{abcd} which we ordinarily regard as the covariant curvature properties!

Let us now justify these claims concerning the dimensional difference between the two curvature tensors. As we saw, the coefficients Γ_{ij}^l are dimensionless pure numbers, and therefore sums of products of these as well as partial derivatives of them are likewise pure numbers. Hence, at any given point, the components of the mixed curvature tensor R_{bcd}^a defined by our Equation (1) must likewise be dimensionless *pure numbers* which are independent of the choice of any unit of length. But, since $R_{abcd} = g_{an} R_{bcd}^n$, the dimensional status of R_{abcd} is exactly the same as that of the metric coefficients g_{an} , which not only have dimension L^2 but also contain the metrical scale factor dictated by the choice of unit of length. Hence R_{abcd} depends on the choice of metric unit, and its L^2 -dimensionality is metrical rather than merely qualitative. It follows that, at any given point, *the components of the (0, 4) curvature tensor each have dimension L^2 , where L is a METRICAL property!*

This dimensional analysis of R_{abcd} was carried out on the basis of the requirement that in every coordinate system the dimensional status of the metric tensor g_{ik} not only be the same, but also that the components g_{ik} contain a metric scale factor dictated by the choice of unit of length. This requirement led to conclusions about the dimensional status of R_{abcd} which hold alike for all components of that tensor and in every coordinate system. But the reader will recall my simple example of the coordinatization of the 2-sphere by means of colatitude ϕ and longitude θ , both of which range over radians, which are pure numbers, so that $d\phi^2$ and $d\theta^2$ MUST be pure numbers. As I emphasized by reference to that example, my refutation of Glymour's purported counterexample does *not* depend on the imposition of the stated requirement that the dimensional status of the metric coefficients be the same in every coordinate system. For it suffices for my argument against Glymour's alleged counterexample that THERE ARE coordinate systems, such as spherical coordinates r, ϕ, θ in Euclidean 3-space, in which at least some of the metric coefficients MUST contain the metrical scale factor dictated by the choice of the unit of length and must be of dimension L^2 . The existence of such coordinate systems has the consequence that there are at least some components of some covariant curvature tensors R_{abcd} which depend on the choice of metric unit and have the dimension L^2 . At least for the sake of simplicity and generality, I shall hereafter continue to employ my general formulations as based on the stated L^2 -dimension-

ality of g_{ik} in every coordinate system without thereby jeopardizing my argument against Glymour's supposed counterexample in the face of someone who does not accept the necessity for such invariance of dimensional status.

The sectional or Riemannian curvature K_N at a point P of a Riemann space with respect to the orientation N determined by two linearly independent contravariant vectors A^i and B^i is given by

$$K_N = \frac{R_{abcd} A^a B^b A^c B^d}{(g_{ac}g_{bd} - g_{ad}g_{bc}) A^a B^b A^c B^d}.$$

But the (0, 4) curvature tensor in the numerator and the (0, 2) metric tensor in the denominator each have dimension L^2 . It follows that the scalar sectional curvature K_N at the point P has the dimension L^{-2} . This sectional curvature K_N is the Gaussian (or total) curvature at the point P of the two-dimensional geodesic surface swept out by geodesics through P which have directions in the two-parameter family of directions $uA^a + vB^a$ at P . Thus, the Gaussian curvature of this geodesic surface at the given point P is the Riemannian (or sectional) curvature of the enveloping n -dimensional Riemann space at P with respect to the given orientation N . And this sectional curvature K_N is of dimension L^{-2} !

This is as it should be, since it follows from Gauss' *theorema egregium* of surface theory which gives the Gaussian curvature K as a function of the metric coefficients g_{ik} and their first and second partial derivatives that K has dimension L^{-2} . We would also expect this on elementary grounds by noting that the Gaussian curvature K of a 2-sphere embedded in Euclidean 3-space is r^{-2} , where r is the length of its radius, and thus depends numerically on the choice of the unit of length.

As will be recalled, on the basis of results *other than* the dimensional analysis of R_{abcd} , I had concluded earlier in this section that "at least with respect to currently known resources for constructing R_{abcd} , the covariant curvature tensor is simply *not defined without a metric tensor*." But now our dimensional analysis of this (0, 4) curvature tensor has shown its dimension to be L^2 , where L is a metrical property. Hence this further result obviates the proviso "at least with respect to currently known resources for constructing R_{abcd} " and entitles me to assert *categorically* that the (0, 4) curvature tensor properties are simply not defined without a metric!

R_{abcd} is no more simply a $(0, 4)$ tensor than a velocity is simply a contravariant vector: As Schouten has stressed (1951, Ch. VI, p. 126), tensors occurring in physics 'are not by any means identical' with objects that merely have the transformation and other purely mathematical properties of tensors. Even though GMD envisions the reduction of all of physics to geometry, the demonstrated L^2 -dimensionality of R_{abcd} is unaffected by this reduction, and thus the latter covariant curvature properties do not accord with the Pythagorean ontological dictum that all is pure number.

Thus, even if — and it is indeed a *big IF* — the NON-METRICAL resources outlined by Glymour sufficed for the construction of a covariant 4th rank tensor that *shares* those particular properties of the $(0, 4)$ Riemann curvature tensor which he lists, I deny flatly on dimensional grounds that he is entitled to the following conclusion of his: "... it would be entirely reasonable to regard the properties he [the putative theoretician] is talking about as what we ordinarily regard as [covariant] curvature properties [R_{abcd}]." Glymour's putative non-metrical $(0, 4)$ tensor does *not* have the METRICAL dimension L^2 , while the $(0, 4)$ curvature tensor does have that metrically-constituted dimension. Therefore, Glymour has failed to show that my inconsistency charge against claims 1–3 is gratuitous or unsound in the context of R_{abcd} -curvature and that the tenability of that charge requires the feasibility of *defining* a *unique* metric by means of the curvature properties. Indeed the demonstrated METRICAL L^2 -dimensionality of the covariant curvature properties shows my inconsistency charge to be true.

Finally, since the sectional curvature K_s has dimension L^{-2} , I see very little more than semantic baptism in Glymour's gambit of *calling* a quantity 'sectional CURVATURE' merely because its values are given by a smooth function from pairs of vectors in every tangent space at every point into the real numbers. Consider an analogy. For a fixed time t , a function from married couples (pairs) into real numbers whose values are the combined ages of the pairs in number of years (at the given time t) cannot be validly claimed to render the joint financial worth in dollars of the respective pairs (at time t) merely because the combined-age values of the former function in number of years are each given by a real number no less than the number of dollars which is the economic worth of any given married pair at time t . Even if, at the time t , the two

numbers which are the values of the respective functions *happen to be the same* for any given married pair, they are each 'impure' (in Carnap's sense) by being respectively a number-of-years and a number-of-dollars.

We are now ready to deal with our question (ii).

(ii) *Question (ii): Were Clifford and Wheeler consistent in their ontology of curvature?* Our question now is whether MUC is inconsistent in the specific context of the assumptions made by Clifford and Wheeler, and if so, whether either Clifford or Wheeler did assert MUC.

W. K. Clifford (1876) gave his Cambridge lecture 'On the Space-Theory of Matter' in 1870 and died in 1879, six years after publishing his English translation of Riemann's Inaugural Dissertation of 1854.¹¹ Riemann elaborated technically on his foundational Inaugural Dissertation in his *Pariserarbeit* of 1861, but the latter was not published until 1876, which was ten years after Riemann's death and only three years before Clifford died (Kline, 1972, p. 889, and Weber, 1953). Furthermore, Clifford's 1870 lecture 'On the Space-Theory of Matter' is both brief and purely verbal rather than mathematical. Hence it is *not* safe to assume that Clifford was familiar with Riemann's 1861 paper when he gave his lecture in 1870 or even by 1875, when he dictated the essay on 'Space' which appeared as chapter II of his posthumously published *The Common Sense of the Exact Sciences* (Clifford, 1950).¹² On the other hand, Riemann's 1854 Inaugural Dissertation was published in 1868, and in the opening paragraph of his 1870 lecture, Clifford wrote (though without giving an explicit reference) that "Riemann has shewn that ... there are different kinds of space of three dimensions; and that we can only find out by experience to which of these kinds the space in which we live belongs." Hence Clifford was almost certainly familiar with Riemann's Inaugural Dissertation when he gave his 1870 lecture. Thus Clifford can be presumed to have been familiar by 1870 with the sectional two-dimensional (Gaussian) curvatures that obtain at any given point with respect to various orientations in n -dimensional Riemann space, since Riemann dealt with these sectional curvatures in his Inaugural Dissertation while *not* mentioning anything corresponding to a fourth rank curvature tensor until his 1861 *Pariserarbeit*.

¹¹ See the posthumously published Clifford (1950, pp. xxx and 247).

¹² The 1875 dictation date of that chapter is given in K. Pearson's Preface to this work, p. LXIII.

But even if Clifford had been aware of the contents of the 1861 *Pariserarbeit* by 1870 or by 1875, there is still ample reason to think that Clifford never had any inkling of the *mixed* Riemann curvature tensor, let alone of the feasibility of its non-metrical construal or of the existence of any kind of non-metrical curvature properties. In the first place, in his 1861 *Pariserarbeit*, Riemann employed only what we now call the *covariant* Riemann curvature tensor, when stating a necessary condition for the isometry of two spaces (Kline, 1972, pp. 894–895 and 1125–1127). Though he can thus be said to have worked with a particular *species* of tensor, Riemann did *not* have the concept of the *genus* tensor, introduced after his and Clifford's death by Ricci and Levi-Civita as an object of rank $r = l + m$, contravariant of rank l and covariant of rank m . Nor did Riemann have knowledge of the tensor calculus as such (Kline, 1972, pp. 1122–1127), and there is no indication anywhere in his work of the concept of contravariance. For these reasons, Morris Kline has expressed the view that "The question as to whether Riemann had any concept of the mixed curvature tensor must, I am quite sure, be answered in the negative. ... there was no tensor analysis in 1861 or even 1870, though the subject had its beginnings in the work of Beltrami, Christoffel and Lipschitz."¹³ Thus, there is good reason to think that even if Clifford did know Riemann's 1861 *Pariserarbeit* when Clifford gave his seminal 1870 GMD lecture, he had no notion then or thereafter of the *mixed* Riemann tensor as a curvature entity, even in the purely metrical construal of that tensor.

In the second place, it was not until 1917–1918, almost four decades after Clifford's death, that mathematicians first dispensed with the Riemann metric when achieving a purely affine generalization of Riemannian metric space such that the mixed curvature tensor can be construed non-metrically (Kline, 1972, pp. 1130–1134 and Weatherburn, 1957). Hence there is every reason to conclude the following: The notion of curvature known to Clifford by 1870 or even by 1875 – the date of his aforementioned chapter on 'Space' – was such that Clifford, no less than Gauss and Riemann, conceived of curvature as a kind of property which is conferred on a space (at any given point) *only* by a metric (tensor). Consequently, in the context of Clifford's construal of curvature, the affirmation of the thesis MUC' would be inconsistent.

¹³ Private communication, quoted with the kind permission of Professor Kline.

Though Clifford's 1870 sketch of his 'Space-Theory of Matter' was wholly non-mathematical, it clearly enunciated the remarkably original thesis that all matter and radiation, including all physical devices which effect spatio-temporal measurements, are constituted out of *empty, curved, metric space*. Said he:

I hold in fact

(1) That small portions of space *are* in fact of a nature analogous to little hills on a surface which is on the average flat; namely, that the ordinary laws of [Euclidean] geometry are not valid in them.

(2) That this property of being curved or distorted is continually being passed on from one portion of space to another after the manner of a wave.

(3) That this variation of the curvature of space is what really happens in that phenomenon which we call the *motion of matter*, whether ponderable or etherial.

(4) That in the physical world nothing else takes place but this variation, subject (possibly) to the law of continuity.

As is evident from our prior analysis, Clifford's own construal of his program of reducing all of physics to geometry required a *metric* both non-trivially – by postulating that space is curved in the pre-1917 sense – and trivially, by postulating that the world's ultimate substance is a metric (presumably Riemannian) space. Moreover, we can now see that the *emptiness* which Clifford attributed to the curved metric space of his monistic vision required the following: Up to a constant positive factor k , both the curvature and the metric of space is 'implicit' in it or intrinsic to it in the sense of at least *not* being *imposed* on the continuous spatial manifold, but of being grounded solely in the very structure of that spatial manifold itself. In particular, as we already saw at the end of §2, any and all kinds of mensurational devices are themselves held to be constituted out of empty *curved* space to begin with. Hence apart from the possible exception of the scale factor k , the non-trivial metric and curvature of space cannot first be imposed on space or first be induced into space by the behavior of any such devices. Any and all bodies or radiation which serve as standards of metric equality qualify as such at best *only epistemically* as means of discovering the metric properties of space. And no entities other than the structure of empty space itself are *ontologically* necessary for endowing space with metric ratios or with such curvature properties as are determined by these ratios. Thus his 1870 lecture committed Clifford to the claim that space is both intrinsically curved *and* intrinsically metric (modulo a scale factor k).

But in his subsequent 1875 essay on 'Space', Clifford (1950) asserted that space *lacks* a (non-trivial) intrinsic metric by declaring:

The measurement of distance is only possible when we have something, say a yard measure or a piece of tape, which we can carry about and which does not alter its length while it is carried about. The measurement is then effected by holding this thing in the place of the distance to be measured, and observing what part of it coincides with this distance... (p. 48). The reader will probably have observed that we have defined length or distance by means of a measure which can be carried about *without changing its length*. But how then is this property of the measure to be tested? We may carry about a yard measure in the form of a stick, to test our tape with, but all we can prove in that way is that the two things are always of the same length when they are in the same place, not that this length is unaltered...

Is it possible, however, that lengths do really change by mere moving about, without our knowing it?

Whoever likes to meditate seriously upon this question will find that it is wholly devoid of meaning (pp. 49–50).

This 1875 statement of Clifford's constitutes a denial of the existence of an intrinsic spatial metric. To see this, note that if space were intrinsically metric to within a constant scale factor k in the sense of Clifford's 1870 Lecture, then it would surely *not* be 'wholly devoid of meaning', as Clifford contends in this 1875 passage, to ask whether the lengths of our familiar measuring standards "do really change by mere moving about, without our knowing it." For, as I explained in Ch. 16, p. 503 and Grünbaum (1970, p. 523), "the supposition of such an unnoticed change would surely have meaning with respect to the presumed implicit metric that confers a curvature on empty space, *provided* that (1) the underlying conception of the empty space manifold has meaning, and (2) there exists a metric implicit to the latter manifold."

We can conclude that the combination of Clifford's 1870 thesis with his cited declaration of 1875 entails commitments which are inconsistent as follows: (1) there is the outright inconsistency that space is and also is not intrinsically metric (to within a scale factor k), and (2) there is the more subtly inconsistent commitment to MUC, whose inconsistency is due to Clifford's construal of curvature as requiring a metric. Incidentally, as we saw in §3(a), Clifford's 1870 assumption that continuous physical space is intrinsically and non-trivially metric (modulo k) had been in direct contradiction with a cardinal explicit tenet of Riemann's Inaugural Dissertation, i.e., with RMH.

Turning to Wheeler, we find the following statement by him in the Foreword to his book *Geometrodynamics* (1962b):

The sources of the curvature of space-time are conceived differently in geometrodynamics and in usual relativity theory. In the older analysis any warping of the Riemannian space-time manifold is due to masses and fields of non-geometric origin. In geometrodynamics by contrast – only those masses and fields are considered which can be regarded as built out of the geometry itself.

... The central ideas of geometrodynamics can be easily summarized.

The past has seen many attempts to describe electrodynamics as one or another aspect of one or another kind of non-Riemannian geometry. Every such attempt at a unified field theory has foundered. Not a change in Einstein's experimentally tested and solidly founded 1916 theory, but a closer look at it by Misner in 1956 (III), gave a way to think of electromagnetism as a property of curved empty space (p. xi).

Unlike Einstein and Schrödinger, who tried *non-metrical* affine approaches to unified field theory, Wheeler cast his theoretical lot with Riemannian *metrical* theories, as he explicitly tells us here. And thus Wheeler's attempt to implement the program of GMD avowedly shared the stated assumptions of Clifford's monistic ontology of empty curved *metric* space. Hence the thesis that continuous space is *devoid* of any non-trivial metric which is intrinsic to within a scale factor k is just as inconsistent with Wheeler's fundamental GMD assumptions as with those of Clifford's 1870 lecture.

Did Wheeler avow that continuous space is devoid of any intrinsic metric in the manner of Clifford's 1875 declaration? I know of no such *outright* declaration by Wheeler. But in Ch. 16, p. 503 and in Grünbaum (1970, p. 523) I cited from a joint paper by Marzke and Wheeler in which these authors clearly seem to conceive of the *metrical ratio* of two intervals of space-time as depending ontologically – and not just epistemically – on the behavior of a metric standard along different routes of transport, rather than as depending only on the very structure of the space-time manifold itself. And I went on to point out there (Ch. 16, p. 504 and 1970, p. 524) that if, in that joint paper, Wheeler had conceived of the specified metrical ratio as depending ontologically only on a metric intrinsic to empty space-time, then "the question of a possible dependence of the ratio of the measures of the two intervals on the route of intercomparison (path of transport) would not even need to arise!" If it is correct, as I think it is, to read this joint paper by Wheeler and Marzke as an implicit denial by Wheeler that empty space-time is intrinsically metric, then Wheeler was being inconsistent no less than Clifford.

Glymour gives no indication at all as to why he does not regard the

textual documentation adduced by me as a cogent basis for the conclusion that Wheeler was thus being inconsistent. Instead, Glymour (1972, p. 338) contents himself with an unavailing autobiographical *obiter dictum*, saying "I am not at all convinced that Wheeler would deny that space has intrinsic metric properties."

We saw that, *apart* from the stated inconsistencies, the basic GMD program of Clifford and Wheeler does require the assumption that continuous physical space is intrinsically (and non-trivially) metric, modulo a scale factor k . Yet, as I have documented in some detail in Ch. 16, §2(c) (i) and in Grünbaum (1970, pp. 515–522), this assumption was explicitly and fundamentally denied by Riemann's enunciation of RMH in his Inaugural Dissertation. Thus, as we already had occasion to emphasize in §3(a), one of the important assumptions required by the GMD ontology of empty space is incompatible with one of the cardinal tenets of Riemann's conception of the foundations of the metric geometry of continuous space.

4. THE TIME-ORIENTABILITY OF SPACE-TIME AND THE 'ARROW' OF TIME

In an unpublished paper engagingly entitled 'Sense and Nonsense About Entropy and Time' and in his (1972), Earman has charged that the level of the philosophic debate on the status of the 'arrow' of time has been low, because of the anachronistic neglect of several alleged facts as follows: (i) Presently available evidence indicates that we live in a general relativistic space-time, so that the status of the 'arrow' of time — which was discussed in chapters 8 and 19 of this book — must derive from the temporal features of that space-time, and (ii) If time does have an 'arrow', then the ontological basis of this arrow is to be sought, in the *first* instance, in the so-called 'time-orientation' (see below) of our actual relativistic space-time, and (iii) Discussions of the relevance of entropy (phenomenological or statistical) to time's arrow which are not either based on the theory of relativity or at least capable of being recast relativistically are not of much interest, and yet many writers who claim such relevance ignore that no generally accepted relativistic thermodynamics or relativistic statistical mechanics exists.

Before entering into the substance of the first two of these allegations,

I am prompted to ask: To whose discredit does it redound that we now have neither an acceptedly satisfactory *relativistic* thermodynamics and statistical mechanics nor an adequate special relativistic analytical mechanics of (electromagnetically) interacting particles? Does this unsatisfactory state of affairs not redound to the discredit of currently employed relativity theory, rather than to the demerit of the import of the best available *non-relativistic* theories for the bearing of irreversible processes on the arrow of time? Is not a similar question posed by the existence of branches of physics and chemistry, such as the quantum theory of atomic and molecular structure, which are still driven to employ the Galilean-Newtonian ST-description rather than the special relativistic one of Einstein-Minkowski, let alone the GTR Riemann-Einstein ST-description? In 1972, Dirac arrived at the sobering judgment that we are further away than ever from a relativistic understanding of quantum phenomena.¹⁴ Given that we employ non-relativistic theories of thermodynamics and statistical mechanics *faute de mieux*, should Earman not thereby be given pause in discounting the value of non-relativistic discussions of time's arrow?

In order to have a more adequate basis for examining Earman's allegations, let us see more specifically what he says. Denoting a general relativistic space-time by ' (M, g) ', where ' M ' represents the appropriate kind of 4-dimensional manifold and ' g ' denotes the appropriate kind of metric (tensor), Earman writes (1972):

Let (M, g) be a space-time. Pick some timelike tangent vector at an arbitrary point $x \in M$. Carry this vector around a closed loop based at x by some method of transport which is continuous and which keeps timelike vectors timelike. If such an operation never leads to time inversion, i.e., never causes the transported vector to point into the lobe of the null cone at x opposite to the original vector, then (M, g) is said to be *temporally orientable* (p. 636).

... Assuming that our space-time is temporally orientable, does it have a particular orientation? (and what precisely is it for a space-time to have or fail to have a time orientation?) If so, where does it come from? These questions are central to a cluster of problems often referred to as "the problem of the direction of time." The most fashionable position has been reductionistic: if it exists, the temporal orientation or direction of time for our world is not an intrinsic feature of space-time itself but is to be analyzed in terms of the nature of certain physical processes. Thus, Reichenbach defined the future direction of time in terms of the entropy increase in 'branch systems'. I want to point out here that there is a simple argument that stands in the way of Reichenbach's position. Reichenbach himself grants that space-time can be treated as a manifold with null-cone structure. But

¹⁴ Discussion remark made by Professor P. A. M. Dirac at the 1972 conference mentioned in footnote 1 of this chapter.

this seems enough to motivate the following Principle of Priority:

(PP) Assuming that space-time is temporally continuous, timelike transport takes precedence over any other method of fixing time direction—that is, if the time sense found by the entropy method (or the like) in two regions of space-time disagrees when compared by means of transport which is continuous and which keeps time continuous timelike, then if one sense is right, the other is wrong.

Combining PP with the following fact:

(F) With Reichenbach's entropy method it is always physically possible and in many cases highly likely (according to statistical mechanics) that there will be disagreement as described in PP.

we can conclude that it is always physically possible and in many cases highly likely that either (a) there is no right or wrong about time direction—talk about what direction is the future and what is past is not meaningful or (b) the entropy method does not generally yield the right result. Reichenbach cannot accept either conclusion, since for him the purified concept of time direction is given by the entropy method (p. 637).

After explaining in more detail what is meant by the 'time-orientability' of a general relativistic space-time, I shall argue the following: Time-orientability is *merely* a *necessary* condition for physically singling out the past and future senses of time uniquely as such. For it assures only the globally consistent *mere oppositeness* of two senses of time by failing utterly to specify *NON-trivially* or non-arbitrarily *which one* of them is the future sense and which one is past! If there is indeed an arrow of time in the physical world, then we expect from an adequate account of it the provision of a physical foundation for singling out the future sense or, alternatively, the past sense—IN ALL, BUT NAME rather than the specification of a property of space-time which, in and of itself, can guarantee merely to single it out IN NAME ONLY. But we shall see that the time-orientability of space-time, and such time-orientation as is vouchsafed by it alone, achieves only the latter and not the former. In particular, it will emerge that time-orientability of a space-time (M, g) contributes no more to the physical foundation for time's arrow than does the existence of two merely opposite senses in the time-continuum of Newtonian particle mechanics, senses whose oppositeness does *not* depend on the existence of irreversible process-types. As I have stressed in chapters 7 and 8 as well as elsewhere (Grünbaum, 1967b, pp. 10-11; 1967a, pp. 150-151, 164-165), this mere oppositeness furnishes the basis for respectively stating *what is meant* physically by reversible and irreversible process-types *to begin with*. Indeed Boltzmann, Eddington, Schrödinger, Reichenbach, and others believed that the characterization

of the future sense of time poses a physical problem, because they were well aware that de facto reversed process-types do presuppose physical relations of temporal betweenness while *failing* to furnish a physical foundation for singling out *one* of the two opposite time senses as either future or past. And it was this awareness which prompted their claim that nomologically irreversible or de facto unreversed process-types first *confer* an 'arrow' on physical time.

Once I have given my reasons below for claiming that the time-orientability of space-time *merely sets the stage* for the physical elucidation of time's arrow, I shall contend that Earman's appeal to entropy *fluctuation* phenomena cannot gainsay this conclusion.

It will be useful to consider briefly the situation in a *Newtonian* world of de facto reversed mechanical process-types. In such a world, the specification of two opposite time senses in the assumedly infinite (non-compact) temporal career of a particle is unproblematic, once we are given the relation of temporal betweenness among the event-members of that career. Newton assumed this relation to have the formal properties of the betweenness on a Euclidean straight line. As may be recalled from ch. 8, pp. 214–216 which was amended in ch. 16, §4, on the basis of this temporal betweenness any two events belonging to the particle's career can then serve to define within that career two time senses which are ordinally opposite to each other. This oppositeness can be reflected by the directions of increase and decrease in a real number coordinatization, while failing to suffice for non-trivially or non-arbitrarily singling out one of the two opposite time senses as the physically future one. Clearly, to just *call* one of the opposite senses the 'future' one is trivial and arbitrary.

Let us denote the direction of increasing time coordinates on this one Newtonian particle career by ' t^+ ' and the direction of decreasing time coordinates by ' t^- '. (This notation allows us to reserve the symbol ' $+t$ ' for positive, as distinct from increasing, values of t , and similarly for ' $-t$ '.)

It is then a simple matter to use Newtonian relations of absolute simultaneity to effect a time coordinatization of all other events so as to *reflect* the *global* existence of two opposite time-senses in the Newtonian space-time: If e_1 and e_2 are events in the career of our primary Newtonian particle, and if some event e_4 is absolutely simultaneous with e_2 while

some other event e_3 is absolutely simultaneous with e_1 , then e_4 has the same time sense with respect to e_4 as e_2 has with respect to e_1 . But note that nothing physically non-trivial or non-arbitrary has thereby been established as to whether that common time-sense is past or future! Surely it is unavailing for physically singling out the future time sense that we may have labeled the latter common time sense ' t^+ ' after physically characterizing it as opposite to the one which we went on to label ' t^- '. In a GTR context, this point is well appreciated by F & P & S. who tell us (1972, p. 72) that in their ST-construction of the null-cone structure, one of their axioms "is intended to express — the possibility of distinguishing between 'future' and 'past' light vectors" but add the caveat: "Only the distinction between two classes matters; we do not introduce any intrinsic difference between future and past here" (my italics).

Let us now characterize time-orientability as a property of a space-time (M, g) with a view to showing that the mere possession of this property fails to furnish an adequate ontological foundation for time's arrow in the physical world. It will then be seen that though failing in this way, the novel importance of time-orientability for our understanding of the arrow of time lies in the following fact: The existence of two merely opposite time senses had never been problematic in pre-GTR physics. Assuming that our actual space-time is a time-orientable (M, g) , this property will emerge from our analysis as merely setting the stage for the existence of the full-blown physical foundations of time's arrow, foundations which go well beyond what is vouchsafed by time-orientability alone.

A few preliminaries will serve to introduce the concept of time-orientability which is, of course, *not* confined to GTR space-times.

We saw in §2(iii) of the present chapter that the infinitesimal null cone structure defined by light rays generates a class $\mathcal{C}$ of conformally related metric tensors g_{ik} , i.e., a class 'ST' of conformally related space-times (M, g) . And it will be recalled from §3(b)(i) that any metric tensor de-

termines a corresponding Christoffel symbol of the second kind $\left\{ \begin{smallmatrix} i \\ j \quad k \end{smallmatrix} \right\}$

or so-called metric connection. Hence each metric tensor g_{ik} belonging to the conformal class $\mathcal{C}$ determines a metric connection, and thereby

the null cone structure generates a class C of metric connections $\left\{ \begin{smallmatrix} i \\ j \quad k \end{smallmatrix} \right\}$, one for each (M, g) in the class ST.

In any space on which a Riemannian metric is defined, we can use the resulting metric connection of that space to give a covariant definition of a concept of the equality of two vectors at different points, or equivalently of the parallel transfer of a vector from one point to another. This definition is a consistent generalization of the familiar Euclidean notion that a vector which remains parallel and equal in length to itself under transport will have the same Cartesian components from point to point. The generalized covariant definition of the constancy of a vector under transport is given by specifying what increments dV^i are required in the components of the vector V^i under transport from the point x^i to a neighboring point $x^i + dx^i$, where the vector becomes $V^i + dV^i$. Specifically, the so-called law of parallel displacement for the vector V^i in the Riemann space of given metric tensor and hence given metric connection $\left\{ \begin{smallmatrix} i \\ j \quad k \end{smallmatrix} \right\}$ is (Adler *et al.*, 1965, p. 50):

$$dV^i = - \left\{ \begin{smallmatrix} i \\ \alpha \quad \beta \end{smallmatrix} \right\} dx^\alpha V^\beta.$$

This generalization of the Euclidean concept of vector equality assures that vectors *preserve their length*, as furnished by the metric of the given Riemann space, under the continuous parallel transport here defined by the law of displacement (Adler *et al.*, 1965, pp. 48–49). Thus, in a space-time (M, g) , a vector which qualifies as time-like at a given point P with respect to g will preserve its 4-dimensional length under parallel transport to another point Q , and hence will also be time-like at Q . But, of course, *not* every type of continuous transport which preserves the time-like character of any such vector qualifies as *parallel* transport.

Let us return to the class ST of conformally related space-times (M, g) and to the class C of metric connections, each of which is generated by the null cone structure. It is now evident from the stated formula for parallel transport that to every member of C , there corresponds a particular rule of parallelly transporting a vector V^i along any given path of the space-time M , one for each member (M, g) of ST. Hence, by generating the classes $\mathcal{C}$ and C , the null cone structure also gives rise to a

well defined class T of rules of parallel transport for vectors associated with M . It is known (Carter, 1971, and unpublished)¹⁵ that either all (M, g) in ST *alike* pass muster as time-orientable or all *alike* fail to qualify as such. Therefore, for the purposes of our impending definition of time-orientability, it is to be understood that the parallel transport employed in that definition may be based on any of the rules belonging to T .

A definition of time-orientability based on the employment of the weaker requirement of a merely continuous mode of vector transport which preserves the time-like character of any such vector in a given (M, g) is known to issue in *the same verdict* concerning the time-orientability of a given $(M, g) \in ST$ as the definition based on parallel transport to be given below. Granted this capability of the definition based on the *weaker* transport requirement, its issuance in a *common* verdict as to time-orientability for *all* members of ST becomes intuitive upon realizing the following: If, as noted on p. 739, the continuous multiplying function of the coordinates (or 'conformal factor') in our earlier definition (ch. 1, p. 19) of a conformal transformation may have only *positive* values, then it is clear that the time-like character of a vector will be preserved at every point of M under conformal transformations of the metric tensor. The definition of time-orientability based on the employment of the weaker requirement of merely continuous transport preserving time-likeness may appear preferable. But I shall give the definition based on the stronger requirement of parallel transport, since it is easier to produce and/or follow proofs of time-orientability employing the latter definition.

Now consider an arbitrary ST -point P and an infinitesimal neighborhood of P . Then the set of all infinitesimal ST -vectors at P is said to be the 'tangent space' T_P at P , and any member of T_P is called a 'tangent vector' at P . As we noted in §2(iii), the light-like tangent vectors at P generate two lobes of an infinitesimal (hyper)cone, which are the boundary of the totality of time-like tangent vectors at P . But this much *allows* that there be *globally* closed time-like paths such that the globally extended light-like world lines *no longer* form two 'halves' of a cone which P separates into disjoint sets of points. Hence whenever we speak of the *two* 'lobes' of the cone, this term is to be understood in the infinitesimal sense, so that P does separate the infinitesimal cone into two disjoint sets.

¹⁵ I am indebted to John Porter for this information.

At each point P of the manifold, there is a set α of time-like tangent vectors which is divided into two subsets: Those that point into or lie inside *one* of the two lobes at P , and those that fall into the other lobe. Intuitively speaking, if two time-like tangent vectors at P which lie inside the *same* lobe are said to have *the same time sense*, the problem of time-orientability of any given (M, g) is the following: If we form the union U_α of all the sets α of time-like tangent vectors respectively associated with the points of the entire manifold, then is the global class U_α of time-like vectors partitioned into *two* equivalence classes with respect to the relation of having the same time sense, on the strength of a non-trivial globally *consistent* extension of that relation to time-like vectors belonging to *different* tangent spaces of M ?

The question of the feasibility of consistently and non-trivially extending an equivalence relation among two entities associated with one and the same point to two such entities respectively associated with *different* points is familiar from our discussion (at the end of §3(a)) of Reichenbach's account of the spatial congruence of solid rods. If two rods are spatially equal under local comparison at essentially the same place P and remain suitably insulated from Reichenbachian 'differential' forces, then we can consistently and non-trivially define them to be congruent not only at P but when they are respectively at *different* places P and Q , because they are also locally congruent at *any* place R *independently of their respective paths of transport from P to R* .

Analogously, we first define having the same time sense for two time-like tangent vectors belonging to the tangent space T_P of an arbitrary point P of a given (M, g) as being constituted by lying inside the same lobe of the null cone at P . Thus, given two time-like tangent vectors **A** and **B** at P , then **A** and **B** will be said to have the same time sense – which we shall write as $\mathbf{A} \equiv \mathbf{B}$ – exactly when **A** and **B** point into the same lobe at P . Now consider another arbitrary point Q of the manifold and a time-like tangent vector **C** at Q . And let us *propose* to define **A** and **C** to stand in the relation $\equiv$ of equivalence of time sense just in case the following is true: **C** at Q stands in the previously defined relation of *local* equivalence of time sense to any time-like vector $\mathbf{A}_{\text{pt}Q}$ at Q which is obtained by the parallel transport ('pt') of **A** from P to Q along some continuous path linking P with Q . Before we ask whether this definition of the extended relation $\equiv$ is consistent, let us issue a *caveat*. The locution 'parallel trans-

port *from* P *to* Q is to be understood mathematically rather than temporally: This locution should *not* be construed as implying that the path of transport linking P with Q must itself be a time-like world line, let alone that Q must be in P 's absolute future. For the path in question is to be any kind of continuous path and hence may well be space-like. It is the transported *vector* which must be time-like all along the path of transport, but not the path itself.

Is our proposed definition of the extended relation $\equiv$ consistent for *every* pair of world points $P, Q \in M$? It will be consistent just in case the *same* verdict as to the truth or falsity of the assertion $\mathbf{A} \equiv \mathbf{C}$ is obtained independently of the path along which the time-like tangent vector $\mathbf{A}$ is transported to obtain a time-like tangent vector $\mathbf{A}_{p,q}$. And if the truth value of $\mathbf{A} \equiv \mathbf{C}$ is thus *path-independent*, then the space-time (M, g) , and indeed any other space-time conformally related to it, will be said to be '*time-orientable*'. We see that time-orientability assures the consistent obtaining of a relation $\equiv$ which generates a partition of the aforementioned global class U_+ of time-like vectors into two equivalence classes or opposite time senses.

Now assume the time-orientability of our actual space-time W , and pick a *particular* time-like tangent vector $\mathbf{A}$ at a world point P . Then call the equivalence class or time sense to which $\mathbf{A}$ belongs U_+^1 , while calling the opposite sense U_+^2 . Are $\mathbf{A}$ and the other time-like vectors in U_+^1 future-directed? Or are they past-directed? Plainly, this question is evaded and not answered by the terminological fiat of, say, now so time-coordinatizing W that the time sense U_+^1 is labeled t^+ and hence dubbed 'future-directed', whereupon the sham baptism is made persuasive by using the upward arrow $\uparrow$ to denote U_+^1 . This gambit can be said to have effected a 'time-orientation' for our assumedly time-orientable W . But is the kind of time-orientation generated by the *nominally* 'future' sense U_+^1 not pathetically helpless to *assure* the singling out of the time sense in which, for example, terrestrial human beings deteriorate biologically? Is it not abundantly clear that the time-orientation effected on the strength of time-orientability alone singles out the future *in name only*?

By contrast, suppose that one of the two opposite time senses is singled out physically *in all but name* by reliance on specified physical features of de facto unreversed or nomologically irreversible process types about which more presently. Such physical singling out does *not* objectively

qualify *that* sense to be '*the*' direction of time, as contrasted with the sense opposite to it, which is correspondingly no less singled out physically in that case. As I emphasized in ch. 8 and elsewhere (Grünbaum, 1967a, pp. 151–152; 1967b, pp. 12–14, and 1971b, Section III), the focus on the *head* of the arrow symbol ' $\uparrow$ ' or ' $\rightarrow$ ', and the misleading effects of the locution '*the flow of time*', have obfuscated the obvious but here overlooked fact that if an event E is physically (not just nominally) in the future of another event E_0 , then E_0 is in E 's past. A set of events ordered by the relation '*later than*' is also ordered by its converse '*earlier than*'. Moreover, as I have argued in detail (Grünbaum, 1971b, Section III), the assertion that '*time flows one way from past to future*' is a mere tautology. Hence even the physical, rather than just nominal singling out of the future time sense does *not* illuminatingly qualify that sense as '*the*' direction or '*the*' orientation of time!

Let us make some remarks here concerning physical, rather than merely nominal, devices for singling out the future sense of time as such, or alternatively, the past sense.

If the so-called CPT-conservation discussed in current particle physics is indeed a law of nature, then the experimental results which have been taken to violate CP-invariance can be held to show that *time*-invariance ('*T*-invariance') is likewise violated by them (Sachs, 1972, p. 595). In particular, assuming these results, a specifiable time-rate r at which long-lived neutral K -mesons decay into *two* π -mesons is then known to violate *T*-invariance as a matter of universal dynamical physical law. Let t^+ be the time sense in which humans deteriorate biologically *and* also the sense with respect to which the 2π -decay mode at the rate r is allowed by these laws. Then the future sense of time can be singled out *physically* as the time-sense in which long-lived neutral K -mesons decay into two π -mesons at the rate r . Suppose that these theoretical conclusions from experiments begun about a decade ago turn out to stand up under further scrutiny. Then the specified K_0 -meson behavior will be the first known case of an *elementary* process-type which is *nomologically* irreversible, among those cases in which we can meaningfully distinguish between law-based (nomological) and merely de facto violations of *T*-invariance (cf. ch. 8, p. 211).

See Append.
§64

Without nomological irreversibility of at least one fundamental process-type, the non-trivial physical singling out of the future time sense requires the invocation of process-types whose physical realization is *de*

facto temporally-asymmetric at least in regard to *statistical* incidence. Under this category, let me now supplement chapters 8 and 19 by taking cognizance of the possibility of the existence of a *denumerable infinity* of branch systems.

Referring to ch. 8, pp. 259–262, assume that the space-time region *R* 'inhabited' by the human species is entropically *typical* as follows: the time-direction of entropy increase in the majority of those branch systems whose careers are contained *in R* is *also* the direction in which most members of ensembles of branch systems *elsewhere* in space-time *typically*, though not universally, increase their entropy. As I emphasized *à propos* of the union of ensembles of statistico-thermodynamic systems treated in ch. 19, § 2, the concept of a *majority* of branch systems and the notion of an entropically *typical* ST-region are each *finitist*. Hence if there are a denumerable infinity of branch systems which individually are in disequilibrium during at least part of their limited careers, then care must be taken to assure the finitist meaning required in the entropic formulations given on pp. 257–260 of ch. 8. This can be done by suitable restrictions on the sets of branch systems to which these finitist formulations are held to pertain. For example, perhaps a satisfactory restriction would be to demand the *random selection* (cf. ch. 19, § 2, p. 656) of sufficiently large finite subsets of branch systems. In any case, given the substantial modifications which I made in Reichenbach's theory of branch systems (see ch. 8, pp. 261–262, and the further details in Grünbaum (1967a, §3)), I see no difficulty in principle for so finitizing the handling of a denumerable infinity of branch systems as to guarantee the following: Granted the assumed physical conditions of the global availability of branch systems in states of disequilibrium, their global statistics of entropy increase unequivocally single out one of the two opposite time senses *in all but name* by assuring an overwhelming statistical preponderance of such increase in one of these two directions! Thus, these entropy statistics, which seem logically quite independent of the aforementioned *T*-invariance violation by K_0 -meson decay, single out the future direction of time, while relativistic time-orientability, and such 'time-orientation' as is vouchsafed by it alone, was seen to be incompetent to do so. Perhaps it is useful to note, by way of mere analogy, that the oppositeness of the left and right halves of any given human being does not single out that person's left half, whereas in almost all people the location of the heart does do so.

Two opposite directions can be defined on an ordinary closed line no less than on an ordinary open line. Specifically, "Cyclic order [$S(ABC)$] is represented geometrically by a class K of points on a directed closed line, with $S(ABC)$ meaning 'the arc running from A through B to C , in the direction of the arrow, is less than one complete circuit'" (Huntington, 1935, p. 4), and "Geometrically speaking, ... serial order may be represented by ... the class of points on a directed straight line (that is, a straight line having a definite 'sense' indicated by an arrow)" (Huntington, 1935, p. 2). By the same token, there may exist two opposite directions in the class of such closed time-like world-lines as are allowed by the GTR infinitesimal null-cone structure, no less than in the class of open time-like world-lines allowed by that structure.¹⁶ Hence it should not occasion surprise that an (M, g) may be time-orientable even though there are globally closed time-like world lines in it, such that a world point P on such a line does *not* separate the global extension of P 's tangent light-cone into two disjoint sets.

What is the significance of Earman's criticism of Reichenbach regarding time-orientability, which we cited at the beginning of the present §4? The reader is now asked to recall my statement in ch. 8, pp. 260–261 and 278–280 of the *import* of my modifications in Reichenbach's theory of branch systems. Is it not clear then that these very modifications (ch. 8, pp. 261–262 and Grünbaum, 1967a, §3) enable my statistico-entropic account of the future time sense to be upheld successfully in the face of the very phenomena of temporally counterdirected entropy increases in branch systems which Earman was able to adduce against Reichenbach's original account? As we saw in ch. 8, pp. 260–261 and 278–280, both Popper and Carnap appealed to entropic fluctuation phenomena in a pre-GTR context to impugn any statistico-thermodynamic account of the arrow of time. And I have explained there and in Grünbaum (1973) why my modified version of Reichenbach's theory of branch systems is immune to their criticism. In the pre-GTR setting in which I formulated my account, time-orientability was unproblematic. But if my account were transposed

¹⁶ This point is erroneously overlooked on p. 203 of ch. 8, where I stated incorrectly that "a closed physical time ... cannot meaningfully be cyclic." My error was inspired by overlooking that in the geometrical interpretation of the postulates for cyclic order, $S(ABC)$ requires that the arc $A \rightarrow B \rightarrow C$ to which it pertains be *less than one complete circuit*, because the assertion $S(ABA)$ is disallowed.

into the context of a time-orientable GTR space-time, it would be similarly immune to criticism by means of Earman's invocation of temporally counterdirected entropy increases in branch systems.

Hence I cannot see that Earman's criticism of Reichenbach can serve to lend any credence at all to the belief espoused by Earman that the time orientability of an (M, g) furnishes an adequate ontological basis for an arrow of time in that (M, g) , as distinct from furnishing a mere necessary condition for it.¹⁷ Instead, the analysis I have given seems to me to show that beyond supplying such a mere necessary condition, such 'time-orientation' of an (M, g) as is vouchsafed by the latter's time-orientability alone singles out the future direction of time *in name only*. And I hope that by exhibiting the role of irreversible or de facto unreversed physical process-types, I have answered the following questions raised by Earman (1972, p. 637): "Assuming that our space-time is temporally orientable, does it have a particular orientation? (and what precisely is it for a space-time to have or fail to have a time orientation?) If so, where does it come from?" For we saw that (1) to 'orient' a given time-orientable space-time is to *make explicit*, via a suitable time-coordinatization, which time-like vectors at different world points have the *same* time-sense, while still leaving unspecified which one of the two opposite time-senses is *non-trivially* future or past, and (2) the latter specification is furnished by law-governed *unreversed* process-types. But as we noted on pp. 796–797, even this non-trivial specification no more renders *the space-time* future-oriented than it renders that 4-manifold past-oriented, although a time-like *vector* will then have one of these two orientations.

¹⁷ Oddly enough, elsewhere Earman recognizes *parenthetically* that 'the distinction between the past and the future in the which-is-which sense' does go beyond 'the existence of a globally consistent time sense', i.e., beyond such time-orientation as is vouchsafed by time-orientability alone. See p. 80 of the Earman paper cited in fn. 20 of the *Appendix*, §13.

BIBLIOGRAPHY

- Adler, R., Bazin, M., and Schiffer, M., *Introduction to General Relativity*, McGraw-Hill, New York 1965.
- Bergmann, P. G., *Introduction to the Theory of Relativity*, Prentice-Hall, New York 1946.
- Carter, B., 'Etiology of Space-Time Manifolds', Cambridge University preprint, 1967.
- Carter, B., 'Causal Structure in Space-Time', *General Relativity and Gravitation I* (1971), 349–391.
- Clifford, W. K., 'On the Space-Theory of Matter', *Proc. Cambridge Phil. Soc.* 2 (1876), 157–158; reprinted in *Mathematical Papers by William Kingdon Clifford* (ed. by R. Tucker), London 1882. Tucker's edition was reissued in 1968 by Chelsea Publishing Co., New York. Clifford's 1870 lecture is also reprinted in *The World of Mathematics* (ed. by J. R. Newman), vol. 1, Simon and Schuster, New York 1956, pp. 568–569.
- Clifford, W. K., *The Common Sense of the Exact Sciences*, Dover Publications, New York 1950.
- Earman, J., 'Some Aspects of General Relativity and Geometrodynamics', *J. Philosophy* 69, (1972), 634–647.
- Ehlers, J., Pirani, A. E., and Schild, A., 'The Geometry of Free Fall and Light Propagation', in *General Relativity* (ed. by L. O'Riartaigh), Clarendon Press, Oxford 1972, pp. 63–84.
- Einstein, A., 'Autobiographical Notes', in *Albert Einstein: Philosopher-Scientist* (ed. by P. A. Schilpp), Tudor Publishing, New York 1949.
- Einstein, A., *The Meaning of Relativity*, 5th edition, Appendix II, Princeton University Press, Princeton 1955.
- Eisenhart, L. P., *Riemannian Geometry*, Princeton University Press, Princeton 1949.
- Glymour, C., 'Physics By Convention', *Philosophy of Science* 39 (1972), 322–340.
- Graves, J. C., *The Conceptual Foundations of Contemporary Relativity Theory*, The MIT Press, Cambridge 1971.
- Graves, J. C., 'Reply to Stein and Earman', *J. Philosophy* 69 (1972), 647–649. This published paper is the truncated form of a longer, unpublished reply to Stein and Earman, which Graves presented orally at the December 27, 1972 Boston meeting of the American Philosophical Association. Citation from the latter is by Graves's kind permission.
- Grünbaum, A., *Philosophical Problems of Space and Time*, Knopf, New York 1963. 1st ed.
- Grünbaum, A. (1967a), 'The Anisotropy of Time', in *The Nature of Time* (ed. by T. Gold), Cornell University Press, New York 1967, pp. 149–186.
- Grünbaum, A. (1967b), *Modern Science and Zeno's Paradoxes*, Wesleyan University Press, Middletown, Conn. 1967. A revised British edition was published by Allen and Unwin in London in 1968.
- Grünbaum, A., *Geometry and Chronometry in Philosophical Perspective*, University of Minnesota Press, Minneapolis 1968.
- Grünbaum, A., 'Reply to Hilary Putnam's 'An Examination of Grünbaum's Philosophy of Geometry'', in *Boston Studies in the Philosophy of Science* (ed. by R. S. Cohen and M. W. Wartofsky), Vol. V, Reidel Publishing Co., Dordrecht 1969, pp. 1–150.
- Grünbaum, A., 'Space, Time and Falsifiability', Part I, *Philosophy of Science* 37 (1970), 469–588. Reprinted in this book as ch. 16.
- Grünbaum, A. (1971a), 'Are Spatial and Temporal Congruence Conventional?', *General Relativity and Gravitation* 2, (1971), 281–284.
- Grünbaum, A. (1971b), 'The Meaning of Time', in *Basic Issues in the Philosophy of Time* (ed. by E. Freeman and W. Sellars), Open Court, LaSalle, Ill., 1971, pp. 195–228.
- Grünbaum, A., 'Karl Popper's Views on the Arrow of Time', in *The Philosophy of Karl Pop-*

- per (Library of Living Philosophers, ed. by P. A. Schilpp, 11: Open Court, LaSalle, Ill., 1973).
- Huntington, E. V., 'Inter-Relations Among the Four Principal Types of Order', *Trans. Am. Math. Soc.* 38 (1935), 1-9.
- Kline, M., *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York 1972.
- Kundt, W. and Hoffman, B., 'Determination of Gravitational Standard Time', in *Recent Developments in General Relativity*, Pergamon Press and The Macmillan Co., New York 1962, pp. 303-306.
- Laugwitz, D., *Differential and Riemannian Geometry*, Academic Press, New York 1965.
- Marzke, R. F., 'The Theory of Measurement in General Relativity', AB senior thesis, Princeton (1959). See also Marzke and Wheeler (1964).
- Marzke, R. F. and Wheeler, J. A., 'Gravitation as Geometry I: The Geometry of Space-Time and the Geometrodynamical Standard Meter', *Gravitation and Relativity* (ed. by H. Y. Chiu and W. F. Hoffmann), W. A. Benjamin, New York 1964.
- Massey, G., 'Is "Congruence" a Peculiar Predicate?', in *Boston Studies in the Philosophy of Science* (ed. by R. S. Cohen and R. C. Buck), Vol. VIII, PSA 1970, Reidel Publishing Co., Dordrecht 1971, pp. 606-615.
- Pirani, F. A. E. and Schild, A., 'Conformal Geometry and the Interpretation of the Weyl Tensor', in *Perspectives in Geometry and General Relativity* (ed. by B. Hoffman), Indiana University Press, Bloomington, Ind., 1966, pp. 291-309.
- Putnam, H., 'An Examination of Grünbaum's Philosophy of Geometry', in *Philosophy of Science* (ed. by B. Baumrin) (The Delaware Seminar, Vol. 2), Interscience Publishers, New York 1963, pp. 205-255.
- Reichenbach, H., *The Philosophy of Space and Time*, Dover Publications, New York 1958.
- Sachs, R. G., 'Time Reversal', *Science* 176 (1972), 587-597.
- Schouten, J. A., *Tensor Analysis for Physicists*, Oxford University Press, London 1951.
- Schouten, J. A., *Ricci-Calculus*, 2nd edition, Springer Verlag, Berlin 1954.
- Stachel, J., 'Space-Time Problems', a review of the Synge *Festschrift* entitled *General Relativity in which the E & P & S paper cited above appears*, *Science* 180 (1973), 292-293.
- Stein, H., 'Graves on the Philosophy of Physics', *Journal of Philosophy* 69 (1972), 621-634.
- Synge, J. L., *Relativity: The Special Theory*, North-Holland Publishing, Amsterdam 1956.
- Synge, J. L., *Relativity: The General Theory*, North-Holland Publishing, Amsterdam 1960.
- Von Laue, M., *Die Relativitätstheorie*, Vol. I, 5th edition, F. Vieweg, Braunschweig 1952.
- Weatherburn, C. E., *Riemannian Geometry and the Tensor Calculus*, Cambridge University Press, Cambridge, England, 1957, 'Historical Note', pp. 176-178.
- Weber, H. (ed.), *The Collected Works of Bernhard Riemann*, 2nd edition, Dover Publications, New York 1953.
- Weyl, H., 'Zur Infinitesimalgeometrie: Einordnung der projektiven und konformen Auffassung', *Nachrichten der Königl. Gesellschaft der Wissenschaften zu Göttingen* (1921), 99-112. Reprinted in 1968, see below.
- Weyl, H., *Space-Time-Matter*, E. P. Dutton, New York 1922. Reprinted in 1950 by Dover Publications, New York.
- Weyl, H. (1923a), *Mathematische Analyse des Raumproblems*, Springer Verlag, Berlin 1923. Reprinted by Wissenschaftliche Buchgesellschaft, Darmstadt 1963. Also available in a collection of monographs entitled *Das Kontinuum und Andere Monographien*, Chelsea Publishing, New York, no date, alien property custodian reprint.
- Weyl, H. (1923b), *Raum-Zeit-Materie*, 5th edition, Springer Verlag, Berlin 1923.
- Weyl, H., *Philosophy of Mathematics and Natural Science*, Princeton University Press, Princeton 1949.

- Weyl, H., 'Zur Infinitesimalgeometrie: Einordnung der projektiven und konformen Auffassung', in *Gesammelte Abhandlungen* (ed. by K. Chandrasekharan), Vol. II, Springer, New York 1968. Reprint of 1921 item above.
- Wheeler, J. A. (1962a), 'Curved Empty Space-Time as the Building Material of the Physical World', in *Logic, Methodology and Philosophy of Science* (Proceedings of the 1960 International Congress) (ed. by E. Nagel, P. Suppes, and A. Tarski), Stanford University Press, Stanford 1962.
- Wheeler, J. A. (1962b), *Geometrodynamics*, Academic Press, New York 1962.
- Willmore, T. G., *Differential Geometry*, Oxford University Press, Oxford 1959.
- Woodhouse, N. M. J., 'The Differentiable and Causal Structures of Space-Time', *J. Math. Physics* **14** (1973), 495–501.

For reasons given in the Preface to this second edition, commentaries will here be made on most of the twenty-two chapters in chapter-by-chapter format. Specifically, the comments below pertain to Chapters 1, 2, 4, 6, Part II (overall), and Chapters 7, 8, 9, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21 and 22. Distinct comments are identified by marginal section (§) numbers.

In doing so, I shall employ the following acronyms or abbreviations: 'GCPP' will refer to my book *Geometry and Chronometry in Philosophical Perspective*; ¹ 'MSZP' will denote my book *Modern Science and Zeno's Paradoxes* in both its first ² and second ³ editions; the term 'Panel article' will be used to speak of any of the six papers in 'A Panel Discussion of Grünbaum's Philosophy of Science', ⁴ and 'Reply to Putnam' will be an abbreviation for the title of my 'Reply to Hilary Putnam's 'An Examination of Grünbaum's Philosophy of Geometry'' ⁵

CHAPTER 1

§1

The Riemannian conception of an intrinsic as opposed to an extrinsic metric, which is extensively invoked in this chapter and indeed in most of Part I, was greatly in need of a more precise explication than is furnished by either the merely intuitive and illustrative statements of it by Riemann and Weyl or by my attempted adumbrations in Chapter 1. Chapter 16, §2 (c) presents a detailed attempt to provide a more precise, systematic and hopefully more coherent formulation than has been available heretofore, although it is probably still beset by inadequacies which others will perhaps be able and interested to remedy. Thus it is essential

¹ University of Minnesota Press, Minneapolis, 1968.

² Wesleyan University Press, Middletown, Conn., 1967.

³ Allen and Unwin, London, 1968.

⁴ *Philosophy of Science* 36 (1969), pp. 331-399; the six contributing authors are G. J. Massey, B. C. van Fraassen, M. G. Evans, R. B. Barrett, G. Wedeking, and P. L. Quinn.

⁵ Published as ch. III of GCPP and, in a slightly different version, in *Boston Studies in the Philosophy of Science* (ed. by R. S. Cohen and M. W. Wartofsky), Vol. V, 1969, pp. 1-150.

that section *B* on Riemann in Chapter 1 be coupled with Chapter 16, §2 (c). Similarly, section *F* on Russell in Chapter 1 is clarified in Chapter 16, §3, pp. 548–550.

The remarks on alternative metrizableability made on p. 15, last full sentence, are to be qualified and supplemented by my essay ‘The Physical Status of the Hypothesis that Everything has Doubled in Size Overnight’ (GCPP, Ch. II, §1) and by §6 of my ‘Reply to Putnam’.

CHAPTER 2

Section A

On pp. 75–77, I gave a *quantitative* argument to show that there is no *a priori* assurance of the existence of at least one time metrization generatable from an admissible time coordinatization and possessing the Newtonian property that the acceleration of every free particle with respect to inertial systems is zero. Mr Emanuel Schorsch has usefully pointed out that this conclusion can be established more readily in the light of the following merely qualitative consideration: In my hypothetical, *contrary-to-fact* example of a particle which is *free* and nonetheless executes simple harmonic, accelerated motion in the Newtonian metric of ephemeris time t , the position and velocity vectors of the particle change their *directions* with time in *every* admissible time coordinatization T . Hence no matter what time metric dT we impose on the Newtonian time continuum by means of the coordinate differentials dT corresponding to any admissible time coordinate transformation $T=f(t)$, the magnitude d^2r/dT^2 of the ensuing new acceleration *cannot* vanish throughout the free particle’s motion.

Hence the incremental utility of my *quantitative* argument on pp. 75–77, which is based on Equations (4) and (5) of pp. 70–71, is just the following: It shows what kind of inadmissible time coordinate transformation $T=f(t)$ is entailed in the given putative case by the demand of Equation (5) on p. 71 that the magnitude d^2r/dT^2 of the ensuing acceleration vanish after all. Schorsch has also generalized my quantitative argument by considering the contrary-to-fact case of free particles whose accelerations with respect to Newtonian (ephemeris) time t *differ* in magnitude at a given instant, instead of being the same, as in the special case of the free particles of my putative example.

§5

Section B

In this section, the theory of a disk rotating with respect to an inertial frame in Minkowski space-time is included under the rubric of the GTR. This inclusion is taxonomically anachronistic. In the more recent construal by relativists of the respective purview of the STR and the GTR, the STR is the theory of *flat* (Minkowski) space-time, while the GTR is the theory of *curved* space-time, where it is to be understood that the 'curvature' ingredient in these two specifications is defined by either form of the fourth rank Riemann curvature tensor, whose meaning is discussed in Chapter 22, §3(b). Thus flat space-time is the special case in which either form of the Riemann curvature tensor is zero. More traditionally, however, the STR was understood more narrowly to be the theory of *inertial* frames in flat space-time as linearly related by the Lorentz transformations. In this narrower sense, accelerated (noninertial) frames like our rotating disk are excluded from the purview of the STR. But on the broader, more recent construal of the scope of the STR, the latter comprises noninertial frames and extends to all reference systems which are covariantly characterized by the fact that the Riemann curvature tensor is zero.

CHAPTER 4

§6

On p. 119, I made the following statement concerning Poincaré's contention that we can always preserve Euclidean geometry in the face of any data obtained from stellar parallax measurements:

if the paths of light rays are geodesics on the customary definition of congruence, as indeed they are in the Schwarzschild procedure cited by Robertson, and if the paths of light rays are found parallaxically to sustain non-Euclidean relations on that metrization, then we need only choose a different definition of congruence such that these same paths will no longer be geodesics and that the geodesics of the newly chosen congruence are Euclideanly related. From the standpoint of synthetic geometry, the latter choice effects a *renaming* of optical and other paths and thus is merely a *recasting of the same factual content in Euclidean language rather than a revision of the extra-linguistic content of optical and other laws*.

This statement calls for two kinds of emendation or *caveat*. In the first place, as pointed out in Chapter 16, pp. 557–558, the locution 'the same factual content' must be understood as elliptical so as to claim equivalence of descriptions only with respect to a specified kind of fact. Secondly,

apart from the first *caveat*, in *synthetic* (as distinct from analytic) Euclidean geometry, light rays will turn out to be non-straight or curved lines under the putative preservation of Euclideanism. But there is no demonstrated automatic assurance that the remetrization which effects this preservation will turn light paths into the *particular kind* of curved line for which a *name* is either available in the already extant vocabulary of synthetic Euclidean geometry or *is definable by means of that vocabulary alone*. Hence unless one resorts to equations which are analytic specifications or 'names' of optical trajectories in a stated kind of coordinate system, one cannot take for granted the feasibility of the 'renaming' of optical paths, which is to effect an equivalent geometric description of them in the manner of Poincaré's renaming of hyperbolic straights as Euclidean semicircles in his so-called 'half-plane' of page 20 in Chapter 1.

Of course, if an analytic specification of optical paths *is* available, then it will codify the desired optical information regardless of whether these paths are held to be straights of a non-Euclidean geometry or members of a synthetically unnamed species of curved lines of a Euclidean geometry.

For a discussion of some related issues pertaining to equivalent descriptions, the reader is referred to a recent paper by Clark Glymour and to references given there.⁶

§7

The treatment on pages 124–125 of alternative geometric interpretations of putative stellar parallax data which constitute interesting live options in the sense of *Duhem's* epistemological holism bears amplification. Take the special case in which stellar light ray triangles are found to have angle sums of less than 180° . This finding permits an inductive latitude in the sense emphasized by Duhem for the following reason: Uncertainty as to the separate validity of the collateral physical laws makes for uncertainty as to what spatial paths are indeed geodesic ones, and this uncertainty, in turn, *allows* that the optical paths are slightly non-geodesic ('curved') in the small *and* markedly *non*-geodesic in the large. Consequently, there is inductive latitude to postulate either of the following two theoretical systems to explain the 'observed fact' that stellar light ray triangles have angle sums of less than 180° :

⁶ C. Glymour, 'Theoretical Realism and Theoretical Equivalence', in *PSA 1970, Boston Studies in the Philosophy of Science* (ed. by R. C. Buck and R. S. Cohen), Vol. VIII, Reidel, Dordrecht, 1971, pp. 275–288.

- (a) G_L : the geometry of the rigid body geodesics is Euclidean
 O_1 : the paths of light rays do *not* coincide with these geodesics, and light ray triangles have area-dependent angle sums of less than 180° .

or

- (b) G_{non-L} : the geodesics of the rigid body congruence are a *non*-Euclidean system of hyperbolic geometry
 O_2 : the paths of light rays *do* coincide with these geodesics, and thus light ray triangles are non-Euclidean ones of hyperbolic geometry.

It will be noted that the optical claims O_1 and O_2 are logically incompatible, although they agree that stellar light ray triangles have angle sums of less than 180° .

§8

On page 146, there is a proposal of an extrication procedure for the sought-after underlying geometry. On pages 193–195 of GCPP, I weakened my initial estimation of the capabilities of this proposed procedure. But I agree with Arthur Fine's subsequent effective criticism in his review of GCPP that even the weakened formulation is essentially unavailing.⁷

§9

The use of physical geometry in Chapter 4 as a crucible in which to examine Pierre Duhem's influential epistemological holism for physics is substantially broadened and revised in Chapter 17. And the reader is referred to the comment on that chapter below.

CHAPTER 6

§10

As noted on page 159, Chapter 6 is devoted to the resolution of Zeno's metrical paradox of extension *as distinct* from his paradoxes of motion, which are first discussed in some detail in Chapter 18 after only cursory consideration on the last page of Chapter 7 (p. 208).

In modern parlance, the legacy of Zeno's metrical paradox of extension can be put as follows. In the modern analytic geometry of spatial and temporal manifolds, a singleton (unit point set or degenerate interval) is assigned the length *zero*, while *non*-degenerate or extended intervals of

⁷ A. Fine, 'Reflections on a Relational Theory of Space', *Synthese* 22 (1971), pp. 476–477. The full text of this review-article covers pp. 448–481.

space or time, which are unions of such singletons, are assigned *positive* lengths (durations). On *physical* (spatial or temporal) grounds, length is governed by additivity requirements which are such that, in the context of the stated assignments of length and duration, the lengths (durations) of all NON-degenerate or extended intervals are paradoxically both zero and positive.

In the ordinary analytic geometry of continuous manifolds, length is countably additive in the following sense: the length of an interval which is divided into a finite or denumerable sequence Σ of nonoverlapping subintervals is equal to the arithmetic sum of the lengths of the members of Σ . (Since there are measure functions on sets which are generalizations of length functions, see also the definition of a countably additive measure in Chapter 16, §2 (a), p. 469). Now, one *cardinal aim* of Chapter 6 is to point out *the reason* for the following conclusion: In the modern mathematical theory *T* of physical space or time, Zeno's metrical paradox of extension is *not* deducible by conjoining *T*'s affirmation of the countable additivity of length (duration) with its assertions that singletons have zero length (duration) *and* that extended intervals are indeed unions of singletons. And, in essence, the reason is that the countable additivity of length (duration) affirmed in *T* is *not* coupled with the assertion that the points of an extended (non-degenerate) interval are only denumerably infinite, i.e., *T* does *not* assert that extended intervals are only countable unions of singletons of zero length. More generally, a major concern of mine in Chapter 6 was to make explicit how the conjunctions of different species of additivity and cardinality in a theory can generate Zeno's metrical paradox of extension.

Some of the infelicities and errors in Chapter 6 were eliminated from the later version of it which appeared as Chapter III in the two editions of my book MSZP. But I believe that despite these emendations, these newer versions of Chapter 6 no less than the original are liable to a criticism contained in the following statement by Massey: "too much emphasis on superdenumerability has masked the remarkable generality of Grünbaum's rebuttal of Zeno's metrical paradox of extension."⁸ Let me quote part of Massey's elaboration of this critical comment and then explain *in what sense* I regard it to be well taken. He writes:

⁸ P.336 of the Panel article by Massey entitled 'Toward a Clarification of Grünbaum's Concept of Intrinsic Metric'.

The soundness and generality of Grünbaum's analysis of Zeno's metrical paradox of extension has been "obfuscated by [the] false impression" that Grünbaum's rebuttal of Zeno depends essentially on "the fact that arithmetic addition is not defined for a super-denumerable infinity of numbers". But imagine that an international congress of mathematicians solemnly endorses the following definition of the sum of a linear continuous series (hereafter, *c*-series) of real numbers: the sum of a *c*-series of real numbers is equal to the least upper bound, if any, on the sums of all the denumerable sequences obtainable by 'deleting' elements of the given *c*-series; if the l.u.b. does not exist, then the sum is undefined. Imagine, too, that this definition proves very useful and fruitful (in presently unspecifiable ways) in measure theory. Now under this definition, the sum of a *c*-series of zeros is zero. Would this mean, then, that Grünbaum's rebuttal is vitiated and that Zeno's paradox is reinstated? Of course not. About the only effect of such a declaration would be to raise the question whether one could define length functions and measure functions on continuous manifolds which were *c*-additive (in the sense that the length of an interval subdivided into a *c*-series of nonoverlapping subintervals is equal to the sum of the lengths of the elements of the *c*-series). Only a moment's reflection is needed to see that the answer to this question is negative. For metrical consistency precludes the *c*-additivity of measures on continuous manifolds, just as it precludes the countable additivity of measures on denumerable dense manifolds and the finite additivity of measures on discrete manifolds. In short, Grünbaum's rebuttal of Zeno depends in no way whatsoever on the historical fact that the mathematical community has not settled upon a fruitful definition of the sum of a *c*-series of real numbers. To believe otherwise is to misinterpret Grünbaum tragically."

My reactions to this very lucid statement fall into several groups as follows:

(1) Let us recall from Chapter 6 not only Luria's reading of Zeno's deduction of metrical paradoxes of extension, which is set forth on page 160, but also the Zenonian additivity arguments given by the mathematical function theorist Paul du Bois-Reymond (p. 160) as well as by the geometer Veronese (p. 173). And let us be mindful of the declaration by P. W. Bridgman, quoted on page 159, who said: "with regard to the paradoxes of Zeno... if I literally thought of a line as consisting of an assemblage of points of zero length and of an interval of time as the sum of moments without duration, paradox would then present itself."

Presumably we are all agreed that it is a straightforward matter to obtain a non-controversially negative answer to the question whether, according to phenomenological thermodynamics, the temperatures of two initially separate liquids are *additive* under the physical operation of pouring the two liquids into the same container. Hence let me ask: Why has there been no perennial discussion of a paradox of the additivity of temperature, whereas Zeno's metrical paradox of extension has seemed

" *Ibid.*, p. 337.

challenging to generations upon generations of mathematically literate thinkers of the caliber of Bridgman? Could it really be that this galaxy of stellar intellects overlooked in the case of length what was so clear for temperature and other metrical attributes? Specifically, did it somehow elude these thinkers that, mathematically speaking, in a discrete manifold, length patently is not finitely additive, and in a denumerable dense manifold, length is not countably additive? Furthermore, if they had been made aware of Massey's conjectured notion of *c*-additivity for a continuous manifold, would it have been *less patent* to them that length is not *c*-additive than that the temperatures of liquids are not finitely additive under the specified physical operation of combining them?

It seems to me that to ask this question is to answer it. I maintain that the whole force of Zeno's dialectic in his metrical paradoxes, and the appeal of these paradoxes to his modern mathematically tutored legates derives from the following fact: Rightly or wrongly – about which more below – the Zenonians either explicitly or tacitly adduced PHYSICAL GROUNDS for *demanding* specified kinds of additivity for length (and duration) in stated types of physical manifolds of space or time! And it was precisely because they regarded the fulfillment of the additivity demands in question as physically mandatory that they used the capability to meet these demands on the part of a proposed mathematical theory of spatial (temporal) structure as a *touchstone* of the adequacy of the given theory. Thus, this physical motivation for the active imposition of specified additivity demands prompted the Zenonians to be adamant about *not* acquiescing in just passively letting any given theory of space or time declare *as a matter of simple arithmetic fact* what sort of additivity is possessed by length (or duration) within its postulational framework without any semblance of paradox.

Hence the Zenonians would not at all have been content to dispose of Zeno's metrical paradox on the basis of the truth of the following statement of Massey's, although they could hardly have hesitated to endorse it: "metrical consistency precludes the *c*-additivity of measures on continuous manifolds, just as it precludes the countable additivity of measures on denumerable dense manifolds and the finite additivity of measures on discrete manifolds." Surely if the Luria reading of Zeno set forth on page 160 is correct, Zeno himself demanded the kind of countable additivity that is meaningful within the mathematical resources of his

time. Hence instead of resolving his metrical paradox for a dense denumerable space by simply denying with Massey that length is countably additive in such a space, Zeno maintained in effect that space cannot be a dense denumerable manifold! By the same token, as we saw in Bridgman's statement, the moral he draws from Zeno's metrical paradox is not all the straightforward mathematical one that there is no paradox because length just is not c -additive in a continuous space or time. Instead, Bridgman discounts the actual structural assertions of the mathematical theory of intervals in physical space and time (and space-time) by saying "if I literally thought of a line as consisting of an assemblage of points of zero length and of an interval of time as the sum of moments without duration, paradox would then present itself." Therefore I believe that one is missing the whole point of Bridgman's remark by pointing to the straightforward mathematical fact that length just is not c -additive in a continuous space or time. But one would indeed be coming to grips with the tacit assumption that prompted Bridgman to deem Zeno's dialectic to be cogent, if one were to challenge the assumption that *on physical grounds* length *must* be c -additive in a continuous space or time.¹⁰ In

¹⁰ Much like Bridgman, the mathematician du Bois-Reymond is prompted, at least in part, by physically motivated demands for the additivity of length to reject the conception of a spatial (or temporal) interval as an aggregate of points such that the intervals of positive length are unions of singletons of zero length. How adamant du Bois-Reymond was on this issue can be gauged from the following statements by him:

"The conception of space as static and unchanging can never generate the notion of a sharply defined, uniform line from a series of points however dense. For after all, points are devoid of size, and hence no matter how dense a series of points may be, it can never become an interval, which always must be regarded as the sum of intervals between points.

The view that the line is a series of points derives once more from the same logical hiatus which we censured twice before. Indeed, the essence of the notion of a series of points forming arbitrarily small intervals between them is not affected by an indefinite continuation of the reduction of the sizes of these intervals. But if we suddenly assume these intervals to be degenerate, then we again arbitrarily and without [adequate] linkage jump to a limit which moreover in this case is not merely unimaginable but outright incoherent or even paradoxical. Hence I reject the generalization of the concept of extension according to which a line is composed of points, a surface composed of lines etc. (*Die Allgemeine Funktionentheorie*, Lauppische Buchhandlung, Tübingen, 1882, Vol. I, p. 66).

... Indeed, one demands the impossible, if a series of points selected from the given points are to determine a point not belonging to the given set. I consider this so unthinkable that I maintain that no intellectual labor can ever extract such a proof for the existence of a limit point from a brain under duress, even if that brain combined Newton's prophetic gifts, Euler's clarity and the devastating power of Gauss' intellect (*ibid.*, pp. 66-67; translation from the German original is mine)."

effect, this is what I did on pages 173–174 of Chapter 6 in reply to Veronese's objections to Cantor's conception of an interval, although Massey's notions of ϵ -sum and ϵ -additivity were then not yet available to me.

(2) Let us distinguish two kinds of countable additivity of length that differ in regard to the character of the subintervals to whose lengths they pertain. The first kind is the countable additivity of length familiar from the ordinary analytic geometry of continuous manifolds and will be referred to as the 'standard' kind. It will be recalled that the standard type asserts countable additivity in the following sense: the length of an interval which is the union of a finite or denumerable sequence Σ of subintervals *no two of which have a common point* is equal to the arithmetic sum of the lengths of the members of Σ . The second kind of countable additivity is one which Massey singled out for consideration (see Ch. 16, p. 477 and n. 14 there). Instead of pertaining to intervals which are nonoverlapping, Massey's type of countable additivity pertains to intervals which are not disjoint but lie 'end-to-end', and can be stated as follows: Consider any interval I which is the union of a finite or denumerable sequence of subintervals such that each member of the sequence has exactly one point in common with its predecessor, if any, and also with its successor, if any, and furthermore is disjoint from all the other members of the sequence. Then the length $L(I)$ is the arithmetic sum of the lengths of these subintervals. To distinguish the latter type from the standard kind of countable additivity, I shall speak of 'Massey countable additivity'.

Although I follow Massey on pages 477 and 489 (Ch. 16) in speaking of his kind of countable additivity as being a 'weak kind', I wish to note that this designation must *not* be taken to mean that Massey additivity is a special case or weaker sense of standard additivity. For intervals lying end-to-end certainly do not qualify as intervals without common points. On the other hand, standard and Massey countable additivity can both be compatibly affirmed within the same theory and indeed are both present in the ordinary analytic geometry of continuous manifolds, the length of a closed interval being the same as that of a half-open or open interval in that theory. Hence the utility of the designation 'weak' for Massey countable additivity lies in conveying *not* a comparison with the 'standard' variety but rather a logical comparison of asserting Massey

additivity, on the one hand, with asserting more strongly both standard and Massey additivity, on the other. Hereafter I shall speak of the stronger assertion of Massey-*cum*-standard countable additivity as the assertion of 'full-blown' countable additivity, when contrasting it with asserting Massey additivity alone.

The difference between full-blown and mere Massey countable additivity of length (duration) is highly relevant, as will presently become clear, to making the following assessment: Has my prior emphasis on the superdenumerability of dense physical space, to the exclusion of its denumerability, been excessive in the face of the need not to run afoul of Zeno's paradox, when meeting the demand for such countable additivity as can be claimed to have *bona fide physical* sanction in a dense physical space? In order to deal with this question, let us first point out that in a *denumerably* dense space, Zenonian metrical inconsistency results from affirming either standard or full-blown countable additivity but *not* from affirming only Massey additivity: As I note on page 489 (Ch. 16): "Massey's *weak* kind of countable additivity... does *not* render the *lengths* of *singletons* additive and hence does not issue in inconsistency with the positive lengths of nondegenerate intervals in either discrete or dense denumerable space." Hence if the imperative to avoid Zeno's metrical paradox while accommodating a physically warranted demand for countable additivity of length is to be capable of effecting a choice between postulating superdenumerably or only denumerably dense intervals, it is crucial whether full-blown or only Massey additivity is held to be mandatory on physical grounds. Therefore, the justification for the claim that, within the genus of dense space, a choice of superdenumerability as against denumerability can indeed be effected via complying with this imperative turns on the legitimacy of a positive answer to the *first* disjunct in the following question, which we now ask: Are there *bona fide physical grounds* for making *full-blown* countable additivity of length (duration) mandatory upon any dense space in geochronometry, or do known physical grounds require the imposition of only the weaker demand for Massey countable additivity?

In some of my earlier writings on Zeno, I asserted the physical legitimacy of requiring *standard* countable additivity. But in others I accepted Massey additivity on physical grounds while stopping short of endorsing the requirement of standard additivity, and I more cautiously merely

called attention to some of the consequences of the latter's incorporation into specified kinds of geo-chronometric theories. Thus, in one of the chapters which I contributed to Wesley Salmon's anthology on Zeno,¹¹ I echoed the demands of the Zenonians to the extent of endorsing standard countable additivity, saying: "it could be reasonably maintained that the *physical relevance* of the metrical concept of length requires its countable additivity" (p. 198). I took the same position in Chapter III of both editions of my MSZP. On the other hand, in Chapter 16 below, which was first published in 1970, I say (p. 477) that, on physical (spatial or temporal) grounds, I assume length to be countably additive in at least Massey's sense. And the treatment of Zeno's metrical paradox of extension which I then go on to give on pages 532–535 of that chapter contains no outright endorsement of standard countable additivity amid pointing out some of the implications of its adoption in certain theoretical contexts.

It will be useful to tackle the question of physical sanction for countable additivity before us by examining the two arguments I gave in MSZP for claiming there that the requirement of standard countable additivity of length is physically warranted. I shall now discuss these two arguments under (A) and (B).¹²

(A) The first argument I offered in MSZP as physical justification for requiring standard countable additivity is given by reference to a linear spatial interval of length 2 units, which is the union of a spatial progression of nonoverlapping subintervals of lengths

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$$

in the given unit. Speaking of this sequence, I wrote (MSZP, Ch. II, §2, pp. 42–43):

Being a progression, our sequence of spatial subintervals has *no last member*. Yet it would be sheer folly mathematically to argue that the absence of a last subinterval in this progression warrants the conclusion that the total interval, which is known to be spatially finite, must paradoxically also be spatially infinite. I mention this folly because we shall need to see why informed people have been driven to commit the corresponding mathematical folly in the case of *time* intervals. But why is it a folly to make the allegation of paradox in the spatial case? Because though denumerably infinite in number, the subintervals are of ever decreasing size such that *all* of them fit into a total finite interval of

¹¹ W. C. Salmon (ed.), *Zeno's Paradoxes*, Bobbs-Merrill, New York, 1970.

¹² When there is a difference in pagination between the first (1967) edition and the revised (1968) edition, I shall distinguish between the two editions by year of publication when giving the appropriate page numbers for each.

length 2 both distributively and collectively, i.e., any one and every one of them fits. That they fit collectively into a finite total interval is evident from the fact that *for every* n , the union of the *first* n of our subintervals in the progression has a length S_n less than 2 given by $S_n = 2 - 1/(2^n - 1)$.

The metrical physical fact that all the ever-shorter subintervals of the progression collectively fit into the total interval of length 2, though the progression is unending, has an important bearing on the additivity of lengths. It shows that we are *justified* in using the standard mathematical limit definition of the arithmetic sum of the infinite sequence of length numbers

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$$

to determine the length of the union of the subintervals, which is equal to the length of the total interval. Thus, we see that the mathematical limit definition of the arithmetic sum of a progression of numbers is indeed appropriate to obtaining the length of the total interval of physical space from the lengths of the members of any particular set of its non-overlapping subintervals which cover the interval up to or including any end-point(s) that it may have. The limit definition of the arithmetic sum of a progression of numbers *cannot*, of course, be invoked *as such* to show that a *finite* interval of physical space can be the union of an unending progression of non-overlapping subintervals. But the use of this limit definition can be justified physically by showing that the lengths of physical intervals are indeed additive in the indicated sense of the arithmetic limit definition.

Does this argument provide adequate physical support for standard no less than for Massey countable additivity or, *as between the two*, only for the latter? I fear that the kind of support furnished by it for the countable additivity of length is such as to sustain the Massey type and not also *all of the commitments* of the standard type. In particular, so far as I can now see, this argument does *not* show physically that if the dense space to which it pertains were only denumerably dense, then the lengths of its singletons would paradoxically have to be additive, thereby ruling out such a space as metrically inconsistent. For note that the subintervals of the example in my argument form a progression, i.e., are ordered *consecutively*. But singletons whose union forms an interval of denumerably dense space and those whose union constitutes an interval of continuously dense space are, of course, arrayed densely rather than consecutively. When I impugned the physical credentials of the Zenonian demand for ϵ -additivity of length in a continuous space, made by Veronese, Bridgman *et al.*, I adduced the denseness of that space on page 174 of Chapter 6 as one of the reasons for giving pause to the advocates of this demand. It would therefore seem that despite the great difference in structure and cardinality between merely denumerably dense and nondenumerably dense intervals, the denseness of the singletons of a denumerable interval should likewise raise doubts about the

physical sanction for demanding full-blown countable additivity over and above Massey additivity. In short, the cited argument does not furnish adequate physical grounds for making full-blown countable additivity of length mandatory upon any dense space over and above Massey additivity.

So far as this argument goes, it now seems to me that, contrary to the Zenonians, in *denumerably* dense space one can grant the *physical intelligibility* rather than the mere metrical fact – of (i) length *not* being countably additive for an interval qua union of a *dense* series of subintervals (singletons) but (ii) length *being* countably additive for an interval qua union of a *progression* of (either non-overlapping or ‘end-to-end’) subintervals.

Before scrutinizing the second of the physical arguments which I gave in MSZP for standard countable additivity, I should relate the above first argument to a clarification of the following statement in Chapter 6, page 164: “Zeno is challenging us to obtain a result differing from zero when determining the length of a finite interval on the basis of the known zero lengths of its degenerate subintervals, each of which has a single point as its only member.” The same statement applies to the position taken by Bridgman in our quotation from him.

Clearly, the formal deducibility of Zeno’s metrical paradox that the length of a non-degenerate interval I is both zero and positive requires the feasibility of obtaining the appropriate kind of *arithmetic sum* of the vanishing lengths of the singletons. This arithmetic sum must be the kind appropriate to the cardinality of the singletons whose union is I , e.g., Massey’s c -sum for a non-denumerably dense interval and the sum of countably many numbers for a denumerably dense interval. Equally clearly, though such arithmetic summability of length-numbers is *necessary* for the deduction of the paradox, the invocation of that arithmetic summability will *not* yield the paradox unless it is ALSO assumed in the theory, on physical grounds, that length is indeed appropriately additive, such as standardly countably additive in denumerably dense space and c -additive in continuous space. Thus, in the above first physical argument, I emphasized (just as I had in the early 1952 paper cited in Chapter 6, page 168n.) that the limit definition of the arithmetic sum of a denumerable sequence of numbers (‘Cauchy sum’) does *not* entail the (standard) countable additivity of length (or duration),

although the assertion of the latter additivity presupposes the Cauchy sum. By the same token, I take the legacy of Zeno's challenge as espoused by Bridgman *not* to rest at all on the fallacy of *inferring* *c*-additivity of length from some arithmetic definition à la Massey of '*c*-sum', or (standard) countable additivity of length from arithmetic Cauchy summability. And assuming the correctness of the Zeno interpretations cited on page 160 of Chapter 6, I presume that if Zeno had been familiarized with the concepts of Cauchy sum and *c*-sum, he would have instantly disavowed these fallacious inferences.

Instead, if he and Bridgman were interviewed together in Elysium, they would endorse, I believe, something like the following elaboration of my aforecited statement on page 164 of Chapter 6. Let us assume that the arithmetic theory employed by geochronometry does provide definitions of *c*-sum and of the sum of countably many numbers such that these definitions are non-trivial generalizations of the ordinary concept of the sum of finitely many numbers. Then physical considerations make it mandatory that length and duration be *c*-additive in continuous space and standardly countably additive in denumerably dense space, so that Zeno's metrical paradox is deducible in either kind of space. Hence the postulate that a stretch of physical space is literally a continuous or denumerably dense aggregate of points is untenable. And if the usual arithmetic theory were to refrain from providing definitions of *c*-sum — as is indeed the case — and of countable sum, then Zeno's paradox is not validly deducible in continuous and denumerably dense space respectively merely for lack of such arithmetic definitions. In Chapter 6, page 173, I put the latter Zenonian point concerning continuous space on behalf of Veronese by saying that Cantor's theory of the interval can claim consistency "only because of the obscurities that obligingly surround the meaning of the arithmetic 'sum' of a superdenumerable infinity of numbers."

In response to this construal of the Zenonian challenge, I undertook the following in my earlier writings: (1) In Ch. 6, I both contested the physical mandatoriness of *c*-additivity of length (p. 174) *and* pointed out that, in any case, the definition of '*c*-sum' which is needed for the formal deduction of the paradox in continuous space is simply not available in the theory (p. 171), (2) I both tacitly (Ch. 6) and explicitly (MSZP) endorsed the physical legitimacy of demanding standard

countable additivity and therefore concluded that dense denumerable space is ruled out by running afoul of Zeno's metrical paradox.

I trust that this clarification removes the liability of my earlier exposition to giving two false impressions to which Massey rightly called attention as such in his Panel article (p. 337). And I now resume the scrutiny of my MSZP endorsement of the physical legitimacy of demanding the standard countable additivity of length.

(B) The second argument which I offered in MSZP as justification for requiring standard countable additivity was the following (1967, pp. 131-132, and 1968, pp. 137-138):

The rejection of *countable* additivity for length and measure would entail incurring the loss of those portions of standard applied mathematics which depend on the presence of countable additivity in the foundations. Thus, for example, one would need to sacrifice some of the mathematics of Fourier series and of the eigenfunctions of quantum mechanics as well as of probability theory and statistics. For countably additive set functions enter into these branches of applied mathematics in one or another form via the Lebesgue integral, the Lebesgue measure, or the Lebesgue-Stieltjes integral.

This argument poses a question as to the precise bearing of physically warranted countable additivity on the rivalry between continuous and denumerably dense space: If we are mindful dimensionally of the ingreience of length and duration in a great many other metrical physical attributes, can the empirical success scored by the usual mathematical physics, which postulates continuous intervals as well as standard and indeed full-blown countable additivity, be held to confer inductive support on full-blown additivity *over and above* Massey additivity? And if so, can it be claimed that the explanatory success of the specified standard mathematical physics redounds to the credit of standard countable additivity to a degree sufficient to militate inductively against a dense *denumerable* geometry, just because the latter can consistently affirm only Massey additivity rather than full-blown countable additivity? For all I know, the answers to both of these questions are affirmative, at least with respect to some one respectable theory of induction or other. But since I cannot claim that this is actually the case, I must now conclude that my second argument in MSZP no more succeeds than the first in adducing adequate physical grounds for making full-blown countable additivity of length mandatory upon any dense space over and above Massey additivity.

It is true that I was primarily concerned in Chapter 6 with noting the *incompatibility* between standard countable additivity and a denumerable geometry. But nonetheless, my failure to have physically vindicated full-blown countable additivity over and above Massey's additivity amid stressing superdenumerability in the avoidance of Zeno's paradox is tantamount to an excessive emphasis on superdenumerability.

GENERAL REMARKS ON PART II

§11

Chapters 7–11 inclusive are listed in the title of Part II of the book as pertaining to the 'topology' of time and space. In his valuable review-article about the first edition, Z. Augustynek has rightly noted that the term 'topology' as used in this context calls for a disclaimer or at least a *caveat*.¹³

Writers such as Carnap and Reichenbach used the term 'topological property' not only in the technical mathematical sense of being invariant under *every* biunique and bicontinuous transformation – about which more on p. 529 of Chapter 16 – but also in the *weaker* philosophical sense of being *non-metrical*. And I followed that tradition of double usage with its attendant risks of being misleading or confusing. For example, both the temporal betweenness in Chapter 7 and the 3-dimensionality of physical space in Chapter 11 are topological properties in the mathematical sense. On the other hand, while the non-metrical dyadic property of being 'later than' among events is weakly topological, that property is not preserved under the particular homeomorphism $t \leftrightarrow -t$ where t is the time coordinate – and hence fails to qualify as topological in the stronger, mathematical sense.

Referring to the title of my Chapter 8, Augustynek calls attention to its construal of the anisotropy of time as the property of non-invariance under the time-transformation $t \leftrightarrow -t$, when possessed by either some basic laws of nature *or* some *de facto* occurring law-governed process-types. And Augustynek then points out that, thus understood, it is at least unclear in what sense this presumed *non-invariance* under $t \leftrightarrow -t$ can qualify as a mathematically topological property of the time continuum: To so qualify, it would itself have to be *invariant* under every

¹³ Z. Augustynek, *Studia Filozoficzne* 45, No. 2, (1966), pp. 202–220 (in Polish). I am most grateful to Professor Augustynek for having translated this review into English for me.

homeomorphism of the time continuum, and in particular also under the topological transformation $t \leftrightarrow -t$ of the time continuum onto itself. Hence in order for the presumed anisotropy of time to be a topological property in the stronger sense, the *non*-invariance of some laws and/or some *de facto* occurring process-types under $t \leftrightarrow -t$ would itself have to be *invariant* under the very same transformation $t \leftrightarrow -t$!

Thus it would seem that if time is anisotropic in the specified sense, its anisotropy cannot meaningfully be construed as a topological property of the time continuum in the *mathematical* sense. Can this anisotropy be understood as topological in the weaker Carnap-Reichenbach sense of being non-metrical?

Certainly not, *if* this sense of 'topological' were held to imply that no time-metric is presupposed in any of the laws or characterizations of *de facto* occurring process-types whose non-invariance under $t \leftrightarrow -t$ makes for the anisotropy of time. For example, insofar as the entropy statistics of the branch systems of Chapters 8 and 19 are temporally asymmetric, the entropic characterization of the anisotropy generated by that temporal asymmetry does involve *time-metrical* specifications in the very concept of the coarse-grained statistical entropy: This presupposition is evident from Chapter 8, pages 237–238 and even more from Chapter 19. Nonetheless, if there are two opposite time directions in the sense discussed in detail in §4 of Chapter 22, then the presumed anisotropy of time *is* weakly topological as follows: Despite the logical ingreience of a time-metric in *physically* singling out uniquely the time sense which we *then call* 'the future direction' – and also in thus singling out the one sense which we then call 'the past direction' – the relations 'later than' and 'earlier than', which are specified thereby, are themselves each non-metrical. But we saw that these relations fail to be topological in the stronger mathematical sense.

We now resume the commentary on the individual chapters.

CHAPTER 7

- §12 The causal theory of time presented in this chapter and applied in Chapter 20, §3, is deficient in a number of respects as follows: (i) It is avowedly restricted to Minkowski space-time and as such is modified, extended or criticized elsewhere by myself and others, (ii) Even within

these confines, it does not deal – as noted on page 687 of chapter 20 – with the import of tachyons¹⁴ or confront the challenge of such causal propagation outside of the line cone¹⁵ as has been thought by some to be allowed by a *compatible* extension of special relativity theory ('meta-relativity'). (iii) It does not develop a causal theory of time within the framework of a causal theory of space-time for flat space-time, let alone for the curved space-times of general relativity. I shall now make some remarks on each of these deficiencies in turn.

§13

(i) Using the acronym 'PPST' to refer to the first edition of this book, I made the following statement in MSZP (p. 58, n.33) to introduce the revision offered there of the strongly modal definition of temporal betweenness given on page 194 of Chapter 7:

The definition of temporal betweenness which I gave for a topologically open time in PPST, pp. 193–195, makes use of the primitive concept of an event being necessary for the k -connectedness of two other events. I am greatly indebted to my colleague Nuel Belnap for having shown me that this primitive concept is dispensable and for having elaborated the alternative definitions that are about to follow – which provide a more satisfactory implementation of the intent of my original construction.

The revision which I then go on to give on pp. 58–60 of MSZP contains a remediable defect which is pointed out, along with its remedy, by G. Massey in his review, forthcoming in *Philosophy of Science*, of B. van Fraassen's book *An Introduction to the Philosophy of Time and Space*.¹⁶ In Chapter VI of that book, van Fraassen gives a discussion of my two formulations (Section 3) before presenting his own 'Systematic

¹⁴ The faster-than-light particles which Gerald Feinberg conjectured within the framework of special relativity and dubbed 'tachyons' in anticipation of their possible discovery have generated a considerable literature. In addition to the references given on p. 687 of chapter 20, let me mention the following pertinent publications in chronological order, some of which contain additional useful literature citations: O. M. P. Bilaniuk and E. C. G. Sudarshan, 'Particles Beyond the Light Barrier', *Physics Today* **22** (May 1969), pp. 43–51; O. M. P. Bilaniuk, *et al.*, 'More About Tachyons', *Physics Today* **22** (Dec. 1969), pp. 47–52; G. Feinberg, 'Particles That Go Faster Than Light', *Scientific American* **222**, (Feb. 1970), pp. 68–77; F. A. E. Pirani, 'Noncausal Behavior of Classical Tachyons', *Phys. Rev.* **D1**, No. 12 (15 June 1970), pp. 3224–3225; M. N. Kreisler, 'Are There Faster-Than-Light Particles?', *Am. Scientist* **61** (1973), pp. 201–208.

¹⁵ Cf. R. Fox, C. G. Kuper, and S. G. Lipson, 'Faster-Than-Light Group Velocities and Causality Violation', *Proc. Roy. Soc. London* **A316** (1970), pp. 515–524; also see J. Earman, 'Implications of Causal Propagation Outside the Null Cone', *Australasian J. Philosophy* **50** (1972), pp. 222–237 and the further references given there.

¹⁶ B. van Fraassen, *An Introduction to the Philosophy of Time and Space*, Random House, New York, 1970.

Exposition of the Causal Theory of Time Order' (Section 4) and its 'Extension to a Theory of Space-Time' (Section 5) for Minkowski space-time. And in his essay 'Temporally Symmetric Causal Relations in Minkowski Space-Time', George Berger¹⁷ comments sympathetically (pp. 60-61) both on van Fraassen's construction and on mine, amid disclaiming (p. 61) to "provide a *relational* theory of space-time based upon causal notions." On the other hand, criticisms of my versions of the causal theory of time and/or of van Fraassen's are offered by Lacey,¹⁸ Smart,¹⁹ and Earman,²⁰ while van Fraassen²¹ has replied to Earman's criticisms in defense of both of us.

§14

Returning to the presentation I give in Chapter 7, let me note first that on p. 180, I define 'separation closure' as the "temporal order exhibiting the formal properties possessed by the order of points on a *closed* (circular) *undirected* line with respect to a tetradic relation of separation," i.e., more accurately, the relation of temporal separation understood *on the model* of the tetradic relation of separation *for points on a circle*. By using the term 'closure', I meant to distinguish between a mere system of separation, which *allows* the *o*-betweenness defined on the same page (p. 180), and the kind of system of separation which, like a circle, does not allow *o*-betweenness. This distinction matters, since a system of *o*-betweenness can be used to define the relation of separation in that system.²² And by speaking of 'separation closure', I meant to confine myself to the species of system of separation which does *not* admit *o*-betweenness. But the formulation I go on to give under (1) on p. 195 of Chapter 7, and perhaps elsewhere in Chapter 7, suggests misleadingly that a system of mere separation is automatically a system of separation-closure, instead of reminding the reader that I am concerned there with the kind of system of separation which rules out *o*-betweenness.

¹⁷ G. Berger, 'Temporally Symmetric Causal Relations in Minkowski Space-Time', *Synthese* 24 (1972), pp. 58-73.

¹⁸ H. M. Lacey, 'The Causal Theory of Time: A Critique of Grünbaum's Version', *Philosophy of Science* 35 (1968), pp. 332-354.

¹⁹ J. J. C. Smart, 'Causal Theories of Time', *The Monist* 53, No. 3 (1969), pp. 385-395.

²⁰ J. Earman, 'Notes on the Causal Theory of Time', *Synthese* 24 (1972), pp. 74-86.

²¹ B. van Fraassen, 'Earman on the Causal Theory of Time', *Synthese* 24 (1972), pp. 87-95.

²² E. V. Huntington, 'Inter-Relations Among the Four Principal Types of Order', *Trans. Am. Math. Soc.* 38 (1935), p. 6, §1.4.

§15

Under the heading of 'Closed Time', I represented Gödel on p. 201 as having claimed that Einstein's theory of gravitation allows closed time-like 'geodesics', although in the footnote I speak more weakly of closed time-like 'world-lines'. But in the publications by Gödel which I cited there, he had not claimed that the closed time-like world lines are geodesics, and had left uninvestigated whether there are perhaps geodesics among them. Then on p. 202, I reported the upshot of a paper by Chandrasekhar and Wright in which the deduction of Gödel's result is impugned. Since then, however, Howard Stein has removed this doubt by showing that Gödel's succinct arguments are quite correct.²³

§16

On p. 203, I said incorrectly that "a closed physical time... cannot meaningfully be cyclic". Both the reason for this error and its correction are stated in Chapter 22, §4, n.16 and in the text to which this n.16 pertains (p. 799).

§17

(ii) To introduce my comment on the issues posed for relativity theory by the hypothesis of super-light causal chains, be they generated by tachyons or otherwise, let me call attention to the appraisal made in Chapter 12, Section F, of Lorentz's appeal to the distinction between true or real times and lengths, on the one hand, and 'local' or 'effective' times and lengths on the other. Lorentz invoked this distinction as a device that would enable him to reconcile the physics of the Lorentz transformations with the *ontology* of the Newtonian space-time structure as embedded in an ether theory. And I call attention to the philosophical objections to this gambit of grafting the ontology of one theory of time order onto a physics of time order which is alien to it: These philosophical objections are set forth both in Chapter 12, Section F, and in Chapter 21.

The ontology of the theory of time-order specified by the Minkowskian coordinate systems of the STR is now as well entrenched as the *Newtonian* one was at the time when Lorentz grafted the *latter* onto the physics of his space-time transformation equations. Let us suppose that perchance microscopic future evidence will bear out the speculation that there is arbitrarily fast tachyonic causal propagation or other such super-light causal propagation after all. Then I ask, how it is not essentially a

²³ H. Stein, 'On the Paradoxical Time-Structures of Gödel', *Philosophy of Science* 37 (1970), pp. 589-601

transposed repetition of Lorentz's philosophical error to insist on adhering to the ontology of the STR time-structure!

Those who insist on the retention of the time-order ingredient in the ontology of Minkowski space-time amid countenancing the existence of such causal chains outside the light cone may then be driven to claim further that with respect to the retained STR theory of time-order, there is *retrocausality* not only in particular reference frames but even invariantly: the effect may temporally precede its cause. In opposition to my claim on p. 189 of Ch. 7, we are told that cause and effect can be distinguished from one another as such (i.e., in the which-is-which sense) by criteria *other than* time order. On this view, our inveterate neglect of this purported fact has created a situation in which "our language is no longer capable of describing a world in which effect may precede its cause," much as the identification of 'north' with 'downstream' in ancient Egypt did not prepare the Egyptians for an encounter with the Euphrates which happens to flow from north to south.²⁴ Thus, speaking of the transmission of information by volition or 'by deliberate action', Paul Fitzgerald says: "there are [*volitional!*] criteria other than temporal order for picking out which of a pair of events causes the other. And these criteria take precedence over temporal priority."²⁵ Fitzgerald then considers kinds of events *R* and *T* which are allegedly such that (1) *R* is *earlier* than *T*, but (2) *R* is identifiable as the reception of a signal while *T* is identifiable as its transmission, i.e., within the pair of causally connected events, *T* is asymmetrically *the* cause of *R*, and *R* is asymmetrically the effect of *T*.

Let it be granted for argument's sake that processes of deliberate action or human(oid) volition are indeed such that in cases of the transmission of information by deliberate action, the following is possible: (a) The respective unique identifications of *T* and *R* as (partial) cause and (partial) effect are *not* viciously logically parasitic on their time order, so that (b) it is logically possible that in the case of one pair of events (*T*₁, *R*₁) which instantiates the event types *T* and *R*, *R*₁ is *earlier* than

²⁴ Cf. P. L. Csonka, 'Advanced Effects in Particle Physics, I', *Phys. Rev.* **180**, No. 5 (1969), pp. 1266-1281, esp. pp. 1266 and 1280.

²⁵ P. Fitzgerald, 'Tachyons, Backwards Causation, and Freedom', in *PSA 1970, Boston Studies in the Philosophy of Science* (ed. by R. C. Buck and R. S. Cohen), Vol. VIII, Reidel, Dordrecht, 1971, p. 421.

T_1 , i.e., T_1 produces R_1 *retrocausally*, while in the case of another pair (T_2, R_2), R_2 is later than T_2 , and the familiar situation of a *tergo* causation obtains.

Even if this be granted for those putative superluminal causal processes which involve deliberate action or (humanoid) volition, it would hardly serve to show that retrocausation is thus logically possible via superluminal causal chains *not* involving any kind of (humanoid) deliberate action or volition. To state my further doubt concerning the retrocausal interpretation of putative superluminal causal chains more fundamentally, consider two causally connected events and denote the one event which is uniquely the (partial) cause of the other by ' C ', while denoting its (partial) effect by ' E ', where C and E generally no longer have the volitional connotations of (humanoid) signaling or communication which adhere to T and R . Thus volition-involving pairs (T, R) are a very special case of pairs (C, E). Then it seems to me that we are entitled to ask the proponents of the meaningful physical possibility of both superluminal retrocausation and subluminal ('tardyonic') *a tergo* causation to furnish us with the following criteria: Physical criteria for the obtaining of the time-order relations of earlier or later between events, and physical — as distinct from volitional — criteria for uniquely identifying C and E respectively as such within any given pair of causally connected events while being fully *compatible* with both *retro* and *a tergo* causation. Let it be noted that when providing the desired criteria for uniquely identifying C and E as such, it is not necessarily illicit for the proponent of both *retro* and *a tergo* causation to avail himself of time-order relations: So long as he can render both *retro* and *a tergo* causation logically compatible with such time order relations as he assumes when stating his criteria for C and E , his logical reliance on time order is *non-vicious*. This distinction between a vicious and a non-vicious reliance on time order will be assumed hereafter.

Perhaps the criteria required for the possibility of both *retro* and *a tergo* causation can be furnished by means of counterfactuals as envisioned, for example, in David K. Lewis's program, which conceives of causation asymmetrically as a special case of *directed* dependence. But this remains to be seen. In this connection, insofar as pertinent, the lessons learned in Chapter 7 from the difficulties which beset Reichenbach's *counterfactual* mark method should be borne in mind as a *caveat*

against vicious logical reliance on time order: A counterfactual analysis of the *asymmetric* causal relation between *C* and *E* must demonstrably avoid the pertinent defects of that Reichenbachian attempt to characterize *C* and *E* asymmetrically. Awareness of these defects can function prophylactically for Lewis's program, even though Reichenbach proposed the mark method *not* in order to allow for the logical possibility of both *retro* and *a tergo* causation but, on the contrary, for the purpose of *defining* the asymmetric *temporal* relations by means of the asymmetric *causal* relations between *C* and *E*. Hence for Reichenbach's mark method *all* logical reliance on time order relations is vicious, whereas the proponent of *retro* and *a tergo* causation can nonviciously avail himself of time order under the stated proviso.

But all this is programmatic. Until the stated criteria for *C* and *E* are provided successfully by the advocates of both *retro* and *a tergo* causation, the physical possibility of non-volitional superluminal *retro-causation* is moot, even if the physical possibility of symmetric relations of superluminal *causal connectibility* among events is taken for granted.

Let us suppose, however, that independently of either the STR or a superluminal extension of it, cause and effect *can* be distinguished from one another without any vicious logical reliance on their relations of time order. I have argued that even then those who are driven to try to accommodate *superluminal* retrocausation in a compatible extension of the STR must tell us why their conception should not stand indicted as committing, in transposed form, the aforementioned Lorentzian philosophical error: Why, I ask, are the relativistic time order relations still ontologically sacrosanct, if there is indeed superluminal causal connectibility among events? I am glad to note that apparently Bilaniuk and Sudarshan²⁶ have implicitly replied that these time order relations are then not ontologically inviolate: They have introduced a universal tachyonic background – perhaps this is tantamount to an ether reference frame whose time qualifies as physically *preferred*? – such that relatively to this special reference frame no superluminal communication is retro-causal.

§18

(iii) Under the heading of a causal theory of GTR space-times and

²⁶ O. M. P. Bilaniuk and E. C. G. Sudarshan, 'More About Tachyons', *Physics Today* 22 (Dec. 1969), pp. 50–52.

times, I can do no more at this time than to call attention to some literature relevant to the pros and cons of its feasibility, in addition to the aforecited exchange between Earman and van Fraassen in *Synthese* (Vol. 24, Nos. 1, 2, 1972).²⁷

In my view, as a first important step, it would be of great interest to be able to give a *causal underpinning* to the GTR space-time structure as elaborated by J. Ehlers, F. A. E. Pirani, and A. Schild.²⁸ This elaboration by these three authors is discussed in other respects in some detail in Chapter 22, §2.

Finally, the merely cursory mention of Zeno's paradoxes of *motion* on pp. 180 and 208 of Chapter 7 is followed up in Chapter 18, which is devoted entirely to issues growing out of these particular paradoxes.

CHAPTER 8

§19 The title of this chapter is 'The Anisotropy of Time'. Hence the reader is asked to recall what I said in §11 above under the heading of an *overall* commentary on Part II of the first edition, which comprises Chapters 7–11 inclusive, concerning the special, non-mathematical sense in which the anisotropy of time can be treated under the heading of the 'topology' of time.

§20 Section A, subsection II, which begins on p. 236, examines the bearing of coarse-grained entropy statistics on the anisotropy of time. These entropy statistics are adduced in Chapter 8 on pp. 238–239 and especially on pp. 257–259, as well as in §4 of Chapter 22. The discussions there of

²⁷ R. Penrose, 'An Analysis of the Structure of Space-Time', Adams Prize Essay, Cambridge University, 1966, partially reprinted in: R. Penrose, 'Structure of Space-Time', in *Battelle Rencontres* (ed. by C. M. DeWitt and J. A. Wheeler), W. A. Benjamin, New York, 1967; E. H. Kronheimer and R. Penrose, 'On the Structure of Causal Spaces', *Proc. Cambridge Phil. Soc.* 63 (1967), p. 481; R. I. Pimenov, *Kinematic Spaces*, Plenum Press, New York, 1970; B. Carter, 'Causal Structure in Space-Time', *General Relativity and Gravitation* 1, No. 4 (1971), pp. 349–391; E. H. Kronheimer, 'Time Ordering and Topology', *General Relativity and Gravitation* 1 (1971), p. 261; R. Penrose, 'Techniques of Differential Topology in Relativity', Monograph published by the Department of Mathematics, University of Pittsburgh, 1971; and S. W. Hawking and G. F. R. Ellis, *The Large-Scale Structure of Space-Time*, Cambridge University Press, Cambridge, 1973, ch. 6.

²⁸ Cf. their 'The Geometry of Free Fall and Light Propagation', in *General Relativity* (Papers in Honor of J. L. Synge), (ed. by J. O'Riada), Clarendon Press, Oxford, 1972, pp. 63–83.

the relevance of the entropy statistics should be coupled with chapter 19. The latter serves as an ontological underpinning for Section A, subsection II of Chapter 8 in the sense of arguing that the coarse-grained entropy of classical statistical mechanics can be construed in scientific realist fashion instead of being a mere anthropomorphism, as has been charged.

§ 22

On p. 235, the analysis of the role of entropic considerations in retrodictive inferences suggests quite misleadingly that past interactions of a rather *detailed kind* can be inferred from the specified entropic information alone. But without premisses going beyond the latter information, the entropically-vouchsafed retrodictive inference of a past interaction of the system with an outside agency cannot yield the conclusion that the outside agency which interacted with the beach system was a stroller.

§ 23

Also on p. 235, and again on p. 255, the erroneous impression is created that Reichenbach's *definition* of a 'branch system' is such as to require that every such system be in a low entropy or disequilibrium state at some time during its career. But as emerges from the account on pp. 258–259 of branch systems which are in equilibrium *throughout* their temporally limited careers, the genus branch system is such that the branch systems discussed on pp. 235 and 255 belong to only one of its species. Thus, as I put it in a 1967 article on the topic of Chapter 8: "Reichenbach points out that there are subsystems which branch off from the wider solar system, galactic system, or from other portions of the universe, remain quasi-closed for a *limited* period of time, and then merge again with the wider system from which they had been separated. And he uses the term 'branch-system' to designate this kind of subsystem."²⁹

§ 24

Since the so-called *fine*-grained entropy of a branch system is a *constant* of the motions of its constituent particles, it throws no light on the temporally asymmetric behavior of such systems and hence receives no attention here.

§ 25

Some of the probability assertions made on p. 239 require buttressing by premisses in addition to those which are explicitly given. The following comments are intended to point out these lacunae.

(1) Let us combine the quasi-ergodic hypothesis of p. 238 with the

²⁹ A. Grünbaum, 'The Anisotropy of Time', in *The Nature of Time* (ed. by T. Gold), Cornell University Press, Ithaca, 1967, p. 161.

statement that many more microstates can realize a homogeneous macrostate than a given non-equilibrium one. These premisses do tell us that the system spends (much) more time in *any one* relatively homogeneous distribution than in *any one* inhomogeneous one. But these two premisses do *not* tell us how the *number* of essentially homogeneous distributions (macrostates) compares to the number of inhomogeneous ones. Thus, taken by themselves, these two premisses allow that the system might *not* spend the bulk of its career near or in equilibrium, as in fact it does. For they still allow that the *total* amount of time spent in the various inhomogeneous distributions altogether might outweigh the time spent near (or in) equilibrium. It is a conclusion from further premisses that the system spends the bulk of any very long finite portion of its career *near* or in equilibrium because the cardinality of the union of the sets of microstates corresponding to the combined latter distributions does far outweigh the total number of microstates corresponding to all the less homogeneous distributions combined.

(2) An entropy curve all of whose dips are of the *same* depth could embody the conclusion that low entropy states are less probable than higher ones, i.e., that the system spends more time in higher states than in lower ones. And a system having such an entropy curve could exist. But in the typical system, there are overwhelmingly more little dips than big ones. As can be understood by reference to a certain randomization assumption made by Boltzmann, for a *typical system*, in the vast majority of its *subequilibrium* entropy states, the system will tend to go back to equilibrium as claimed by Tolman³⁰ and by Reichenbach,³¹ who tell us that large entropy fluctuations are much more improbable than smaller ones.

§ 26

As pp. 254 and 259–262 of Chapter 8 explain, and as is stressed again in Chapter 22, §4, my appreciation of the contribution made by Reichenbach's theory of branch systems is not tantamount to my regarding his formulation of it to be sound: *Non dico eum ab omni naevo vindicari posse*. But in Chapter 8 my proposed altered version of Reichenbach's conception contains a minor error on pp. 256–257 as to the number of

³⁰ R. C. Tolman, *The Principles of Statistical Mechanics*, Oxford University Press, Oxford, 1938, pp. 156–157.

³¹ H. Reichenbach, *The Direction of Time*, University of California Press, Berkeley, 1956, p. 111.

microstates which underlie the totality of occurrences of a given kind of entropy state in a particular system: This numerical error is corrected in n. 4 of Chapter 19, pp. 664–665, and the basis for this correction is developed in the text of Chapter 19, §2 which issues in this footnote. H. Krips³² has offered an alternative to my account of the role of boundary conditions in the entropy statistics of branch systems: Krips's alternative involves a boundary condition at one time on a finite system, whereas my account involves spatially and temporally much more far-flung boundary conditions.

§27 The 1963 treatment of Popper's views given in Section B of Chapter 8 should be supplemented by my updated critical analysis of Popper's ideas on the arrow of time as elaborated in more recent publications of his.³³

§28 The account I give of time's arrow in Chapters 8 and 19 is non-relativistic. Hence the problem of distinguishing between the past and future directions of time in the sense of providing (non-trivial) physical criteria as to *which-is-which* is taken up again specifically vis-à-vis the GTR in Chapter 22, §4, which is devoted to 'The Time-Orientability of Space-Time and the 'Arrow' of Time'.

§29 Finally, I should call attention to some further literature on this problem. H. Mehlberg has included a new treatment of the anisotropy of time in a substantial paper devoted to significant changes which quantum physics has made necessary in our understanding of physical time.³⁴ This paper is a partial account of the results presented in his forthcoming book *Time, Causality and the Quantum Theory*, which is to appear as a volume of *The Boston Studies in the Philosophy of Science*. His earlier views on time's arrow, as well as J. J. C. Smart's and mine were criticized by J. Earman.³⁵ But rebuttals to these criticisms have

³² H. Krips, 'The Master Equation and the Arrow of Time', *Il Nuovo Cimento* **3B** (1971), pp. 165 and 168–169. For a related point of view, see B. Gal-Or, 'The Crisis About the Origin of Irreversibility and Time Anisotropy', *Science* **176**, No. 4030 (1972), pp. 11–17.

³³ See A. Grünbaum, 'Karl Popper's Views on the Arrow of Time', in *The Philosophy of Karl Popper* (ed. by P. A. Schilpp), (Library of Living Philosophers), Open Court, LaSalle, Ill., 1973.

³⁴ H. Mehlberg, 'Philosophical Aspects of Physical Time', *The Monist* **53** (1969), pp. 340–384.

³⁵ J. Earman, 'Irreversibility and Temporal Asymmetry', *J. Philosophy* **64** (1967), pp. 543–549, and 'The Anisotropy of Time', *Australasian J. Philosophy* **47** (1969), pp. 273–295.

been published by both H. Krips³⁶ and G. Berger.³⁷ And in the aforementioned §4 of Chapter 22, I have set forth my reasons for being quite unconvinced by Earman's later attempt³⁸ to find an adequate physical basis for time's arrow without reliance on irreversible process-types.

A number of essays dealing with the anisotropy of time, all of which are reprinted from the 1970 and 1971 volumes of *Studium Generale* and some of which discuss my views, are available in an anthology *The Study of Time*.³⁹

CHAPTER 9

- §30 Problems of statistical explanation and prediction relevant to a needed *reappraisal* of the symmetry thesis concerning explanation and prediction are discussed by Hempel,⁴⁰ Salmon,⁴¹ Rescher,⁴² and Coffa.⁴³

CHAPTER 10

- §31 The conception of the mind-dependent status of temporal becoming advocated in this chapter is elaborated considerably further in Chapter I of my MSZP. And I have published two successive revisions of this chapter,⁴⁴ as well as a reply to cybernetic arguments against my contentions.⁴⁵ For discussions of temporal becoming not cited in these publica-

³⁶ H. Krips, 'The Asymmetry of Time', *Australasian J. Philosophy* 49 (1971), pp. 204-210.

³⁷ G. Berger, 'Earman on Temporal Anisotropy', *J. Philosophy* 67 (1971), pp. 132-137.

³⁸ J. Earman, 'An Attempt to Add Some Direction to "The Problem of the Direction of Time"', forthcoming in *Philosophy of Science*.

³⁹ J. T. Fraser, F. C. Haber, and G. H. Müller (eds.), *The Study of Time*, Springer, New York, 1972.

⁴⁰ C. G. Hempel, *Aspects of Scientific Explanation*, The Free Press, New York, 1965, Part IV.

⁴¹ W. C. Salmon, *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh, 1970, pp. 21-22, 53, 56, 79, 97 and 102, and 'Theoretical Explanation', forthcoming in the *Proceedings of the September 1973 Bristol Conference*.

⁴² N. Rescher, *Scientific Explanation*, The Free Press, New York, 1970, Part II.

⁴³ J. A. Coffa, *Foundations of Inductive Explanation*, University of Pittsburgh Doctoral Dissertation (Philosophy), 1973.

⁴⁴ A. Grünbaum, 'The Meaning of Time', in *Essays in Honor of Carl G. Hempel* (ed. by N. Rescher), Reidel, Dordrecht, 1970, pp. 147-177; and in *Basic Issues in the Philosophy of Time* (ed. by E. Freeman and W. Sellars), Open Court, LaSalle, Ill., 1971, pp. 195-228.

⁴⁵ A. Grünbaum, 'Are Physical Events Themselves Transiently Past, Present and Future? A Reply to H. A. C. Dobbs', *British J. Philosophy of Science* 20 (1969), pp. 145-153.

tions, including some that take issue with my views, I refer to the anthology *The Study of Time* which was cited in n. 39. Those of my papers cited in n. 8 on p. 322 of Chapter 10 which deal only with the free will problem are superseded by my 1972 essay 'Free Will and Laws of Human Behavior'.⁴⁶

CHAPTER 11

The literature cited on pp. 332–333 regarding the question 'Why is Physical Space Three-Dimensional?' should be supplemented by the writings of Whitrow, who argues that the development of higher forms of life would be impossible in a space of fewer than three dimensions,⁴⁷ and of Büchel, who contends after giving a useful survey that "A variety of arguments point to the necessity of fixing the number of dimensions of space at $n=3$."⁴⁸ Furthermore, A. M. and V. M. Mostepanenko have claimed that "The topology of our space including three-dimensionality apparently is conditioned by fundamental microprocesses,"⁴⁹ while Faris has noted that "the existence of physical quantities (such as the magnetic induction vector **B**) that may be represented as axial vectors leads to the conclusion that space is three dimensional."⁵⁰ Finally, Jammer has offered suggestions as to how the dimensionality of *space-time* might be derived from a generalized version of the Mach Principle.⁵¹

⁴⁶ A. Grünbaum, 'Free Will and Laws of Human Behavior', in *New Readings in Philosophical Analysis* (ed. by H. Feigl, K. Lehrer, and W. Sellars), Appleton-Century-Crofts, New York, 1972, pp. 605–627. A slightly different earlier version appeared under the same title in the *Am. Phil. Quart.* 8 (1971), pp. 299–317.

⁴⁷ G. J. Whitrow, *The Structure and Evolution of the Universe*, Harper and Row, New York, 1959. See also his 'Why Physical Space has Three Dimensions', *British J. Philosophy of Science* 6 (1955), pp. 13–31.

⁴⁸ W. Büchel, 'Warum hat der Raum drei Dimensionen?', *Physikalische Blätter* 19, No. 12 (Dec. 1963), pp. 547–549. Also reprinted in Appendix I of W. Büchel, *Philosophische Probleme der Physik*, Herder, Freiburg, Germany, 1965. This citation is from p. 1224 of I. M. Freeman's translation and adaptation of Büchel's article, as published in the *Am. J. Phys.* 37 (1969), pp. 1222–1224.

⁴⁹ English abstract of a paper 'Why Our Space Has Three Dimensions', published in Russian during the last few years by these two authors from Leningrad, according to a private communication from the authors.

⁵⁰ J. J. Faris, 'Comment on "Why is Space Three-Dimensional?"', *Am. J. Phys.* 38 (1970), p. 1265.

⁵¹ M. Jammer, Koerner Lecture (unpublished).

and Earman has posed the issue of whether the question as to the tridimensionality of space is a sensible one to begin with.⁵²

CHAPTER 12

§ 33 Section B on Einstein's conception of simultaneity is substantially extended in Chapter 20, which is centrally devoted to simultaneity. Newtonian as well as special relativistic. And that later chapter corrects or refines some items in both Section B and Section C of Chapter 12 either implicitly or explicitly. Thus, p. 370 of Chapter 12 had addressed itself to the logical situation which would have resulted *if* in 1905, contrary to assumption (i) of p. 346, material clocks had been thought to behave *Newtonianly* under transport while light had then been thought to have the limiting role codified by assumption (ii) of pp. 350 and 369–372. It is pointed out on pp. 692–693 of Chapter 20 that then one of the logical consequences would have been the following: The members of the 1905 scientific community would have been fully entitled to reject Einstein's *relativity of simultaneity* but *not*, as I have it on p. 370 of Chapter 12, "his *conventionalist* conception of *one-way* transit times and velocities."

§ 34 As part of a lengthy 1969 article on the genesis of the STR, Gerald Holton⁵³ takes issue with some of the *historical* facets of Section C. As I noted there, Polanyi had marshalled the history of the STR in support of his view of scientific discovery and knowledge. And as against Polanyi's thesis, I had claimed on p. 382 that the history of the STR is consonant with

an account of both the genesis and justification of scientific theory which has the following two characteristics: first, while being *empiricist*, it is broad enough to accommodate the *valid core* of the Kantian emphasis on the *active*, creative role of the scientific imagination in the postulational elaboration of hypothetico-deductive theories, and second, by being empiricist, it avoids the notorious aprioristic pitfalls of classical rationalism, pitfalls on whose brink Polanyi hovers in the protective twilight of his thesis of the unspecifiability of fallible clues.

⁵² J. Earman, 'Is 'Why is Space Three-Dimensional?' A Sensible Question?', (unpublished, so far as I know).

⁵³ G. Holton, 'Einstein, Michelson, And the 'Crucial' Experiment', *Isis* 60, Part 2, No. 202 (1969), pp. 133–197. This essay is reprinted in Holton's collection of essays *Thematic Origins of Scientific Thought*, Harvard University Press, Cambridge, 1973, II, Chapter 9, pp. 261–352.

On that same page, I immediately went on to ask rhetorically:

Can the history of the STR as conjectured by Polanyi be validly adduced here to prove any more than the untenability of a *crudely* empiricist Machian or Aristotelian-Thomist, abstractionist account of theory-construction as a mere codification of the results of experiments? And is the untenability of the latter kind of account not fully recognized by any empiricist conception of scientific knowledge which incorporates the lesson of Kant? Does such an empiricist conception not allow for the difference between knowing of an experimental result in a narrow sense, on the one hand, and speculatively assigning a wider significance to it, on the other?

That the answer to the last two of these questions is indeed 'yes' is indicated by a single chapter title in a book by an empiricist writer whose views are presumably anathema to Polanyi: Chapter Six of Reichenbach's *Rise of Scientific Philosophy* is entitled 'The Two-fold Nature of Classical Physics: Its Empirical and Its Rational Aspect'.

Ignoring these statements of mine about the value of 'the lesson of Kant', Holton cites my immediately preceding pp. 380–381 in support of the following claims, made by him either directly or indirectly:⁵⁴ (i) As a follower of Reichenbach, on whom the lesson of Kant is alleged to have been lost, I am purported to be animated by 'experimenticism' of which Holton says: "There exists a view of science at the extreme edge of the time-honored tradition of empiricism that will here be called *experimenticism*. It is best recognized by the unquestioned priority assigned to experiments and experimental data in the analysis of how scientists do their own work and how their work is incorporated into the public enterprise of science."⁵⁵ Qua alleged 'experimenticist', I entertain "The desire to see a theory as a logical structure, built upon an empirical basis and capable of verification or falsification by more experiment."⁵⁶ (ii) Being driven by this *normative* view as to the epistemic requirements for a good scientific theory – which presumably earned me the label 'the epistemologist A. Grünbaum' from Holton⁵⁷ – I am said to have produced "implicit 'history' ":⁵⁸ My purported belief that these normative requirements must actually have been fully met in the genesis of the STR led me to allegedly otherwise unfounded or dubious historical imputations and conjectures, particularly in regard to the genetic role of the Michelson-Morley experiments.

⁵⁴ *Ibid.*, pp. 150–151 and p. 165.

⁵⁵ *Ibid.*, p. 147.

⁵⁶ *Ibid.*, p. 151.

⁵⁷ *Ibid.*, p. 180.

⁵⁸ *Ibid.*, p. 151.

In a section entitled 'Explicit History', Holton proceeds to try to show, among other things, how his anti-'experimentalist' conception of the genesis of the STR, which is akin to Polanyi's, can accommodate the first-hand information provided in the *first* installment of R. S. Shankland's 'Conversations with Albert Einstein'.⁵⁹ When Holton wrote his 1969 paper, only this first installment of Shankland's report had been published. But Shankland has since been prompted – possibly even partly by the desire to contest Holton's historical interpretation? – to amplify his first report of 1963 a decade later by a second installment which appeared in July 1973. Says Shankland:

In 1963 I reported on five visits with Prof. Einstein in Princeton during the period 1950–1954. These discussions were published almost verbatim with but little comment by me. They have since been referred to in several articles on the history of physics, so it now seems appropriate to supplement the first publication by a more complete discussion of certain statements made to me by Prof. Einstein. I would like especially to record my impressions of Einstein's views on the role of the experiments of Michelson and Morley and others in the development of the theory of relativity and also to comment more fully on his attitude on quantum mechanics.⁶⁰

From the epistemological point of view, the fundamental issue of scientific rationality in the genesis of the STR is whether some kind or other of pre-relativistically anomalous experimental evidence was genetically relevant, and not whether the Michelson-Morley experiment in particular was so relevant. But some of the discussion of whether the STR was developed rationally has focused on the role of that experiment: On this point, Holton adduces 'direct evidence'⁶¹ from 'Explicit History'.⁶² to invalidate my supposedly aprioristic 'implicit history'. Let me therefore cite the pertinent upshot of the *second* installment of Shankland's report and leave it to the reader to decide to what extent it vindicates me historically vis-à-vis Holton's objections. Shankland writes:

The writer is convinced on the basis of his discussion with Prof. Einstein that he certainly knew about the Michelson-Morley result before 1905 and also was conversant with a wide range of both the experimental and theoretical work in the physics literature bearing on the 'aether problem' before 1905. However, the exact relation which any specific

⁵⁹ *Am. J. Phys.* **31** (1963), pp. 47–57. This important documentary paper appeared in January 1963, too late to be discussed in the first edition. But account is taken of it in Chapter 21.

⁶⁰ R. S. Shankland, 'Conversations with Albert Einstein, II', *Am. J. Phys.* **47** (1973), p. 895.

⁶¹ G. Holton, *op. cit.*, p. 151.

⁶² *Ibid.*, p. 152.

experiment or the theoretical work of other physicists bore to his own creations is not possible to determine with certainty; even Einstein himself was not sure.

... It should be noted that in 1905 it was not the practice to give specific references in published papers as it is today. Many important papers gave no references whatever, so the fact that Einstein's 1905 paper (which has no references) makes no explicit mention of the Michelson-Morley experiment is not in the least unusual. To what degree and at what stage of his activities he was directly influenced by the work of others, it is now impossible to determine. But I became convinced that his interest in the Michelson-Morley experiment (both those of 1886 and 1887) had existed before 1905.⁶³

Shankland's first-hand report likewise calls for a reappraisal of Gary Gutting's adjudication of my dispute with Holton.⁶⁴ Gutting was concerned to 'provide a basis for reconciling' what 'are valid claims made by both Holton and Grünbaum' in a paper whose abstract reads as follows:

This paper discusses the controversy between philosophers of science (e.g. Grünbaum) and historians of science (e.g. Holton) regarding Einstein's discovery of STR. Although Holton is surely correct on the historical point that experimental results (especially the Michelson-Morley experiment) had little influence on Einstein's development of STR, this fact is not sufficient to establish his (and Polanyi's) claim that major scientific discoveries are primarily matters of private, nonspecifiable insights into physical reality. It is possible that Einstein's work was based primarily on non-empirical but nonetheless publicly discussable, objective considerations. And a more comprehensive survey of the discovery of STR shows that this was indeed the case and thus excludes STR as a supporting instance of Holton's and Polanyi's assertions of the primacy of 'private science'.⁶⁵

§35 Holton⁶⁶ also makes a point of the fact that in a 1964 paper in *Science*, which is reprinted above as Chapter 21, I had endorsed Hempel's doubts that the methodologically important attribute of a hypothesis which I had called 'being logically *ad hoc*' in Chapter 12, pp. 387–388 can be defined in purely logical terms. In particular, as Chapter 21 shows, I then stressed Hempel's point that in this context the requisite concept of independent testability involves the *pragmatic* notion of a significantly or interestingly different observational consequence. But I retained the distinction between an auxiliary hypothesis which is in fact *ad hoc* in the admittedly pragmaticized logical sense, on the one hand, and a hypothesis which is not thus *ad hoc* but is only *mistakenly believed* to be,

⁶³ R. S. Shankland, 'Conversations with Albert Einstein, II', *op. cit.* (1973), p. 898.

⁶⁴ G. Gutting, 'Einstein's Discovery of Special Relativity', *Philosophy of Science* 39 (1972), pp. 51–68.

⁶⁵ *Ibid.*, p. 51.

⁶⁶ G. Holton, *op. cit.*, p. 180.

thereby qualifying as 'psychologically *ad hoc*', on the other: Clearly, if there is agreement as to what observational consequences pragmatically count as significantly or interestingly different, then it is a fact of logic rather than of psychological belief whether a given auxiliary hypothesis is or is not contextually *ad hoc* in the pragmaticized logical or 'systemic' sense. In short, my recognition of the presence of a strongly pragmatic component in the attribute of being systemically *ad hoc* does *not* obliterate or even blur the distinction which I drew on pp. 387-388 between being systemically *ad hoc* and being psychologically so, a distinction reiterated on p. 718 of Ch. 21.

Hence it is at least misleading if not a misunderstanding on Holton's part to cite my 1964 paper in *Science*⁶⁷ as support for saying "the epistemologist A. Grünbaum's recent analysis that had confidently announced clear distinctions between 'logically *ad hoc*' and 'psychologically *ad hoc*' soon had to be withdrawn by its author, as the result of the demonstration by Hempel that there were serious inadequacies in the work."⁶⁸ It would be a misunderstanding here to blur the distinction between systemic and psychological *ad hoc*ness by conflating such psychological elements as are ingredient in the pragmatic component of being systemically *ad hoc*, on the one hand, with the psychological nature of a *mistaken belief about* whether an auxiliary is systemically *ad hoc*, on the other.

Moreover, Holton simply flies in the face of an explicit contrary statement of mine when he suggests that I "dismissed as 'merely psychological' " the 'clearly derogatory sense' in which 'Einstein and others' called the Lorentz-Fitzgerald contraction hypothesis *ad hoc*."⁶⁹ For in my 1964 *Science* article (p. 1409; see also Ch. 21, pp. 716-717), I had commented on Wertheimer's personal report of Einstein's reaction to the $L - F$ contraction hypothesis by saying:

I take Einstein's rejection of the contraction hypothesis as *ad hoc* not only to allow but to assert that this hypothesis is indeed 'independently testable', that is, testable by experiments of a kind different from the Michelson-Morley experiment. For I construe Einstein's pejorative usage of the term '*ad hoc*' to refer to the posited fact that though independently testable, the contraction hypothesis would fail to secure subsequent independent experimental confirmation *as against the claims of a new rival theory*.

⁶⁷ *Ibid.*, p. 180 and n. 137.

⁶⁸ *Idem.*

⁶⁹ *Idem.*

We shall see after some analysis that if Einstein did consider the contraction hypothesis as *ad hoc* in this sense, he was quite right. The analysis that vindicates Einstein will also completely *refute* the standard textbook indictment of the contraction hypothesis, which charges this hypothesis with being *ad hoc* in the quite different sense of not being testable independently of the Michelson-Morley experiment.

This reading of Wertheimer's report accords with Einstein's own statement of 1907.⁷⁰

Holton gives his own analysis of the status of the $L-F$ contraction hypothesis vis-à-vis "the operative sense of *ad hoc* hypotheses that does exist among scientists."⁷¹ Since Holton had complained that normative as distinct from descriptive epistemology can generate false history, let me note Elie Zahar's appraisal of Holton's treatment of the *ad hoc* issue. Zahar argues that "Holton is wrong... and that his (rather Polanyiite) methodology misleads him into false history."⁷²

Holton has published a valuable sequel to his 1969 study in an essay on Mach and Einstein.⁷³

§ 36

Turning to Section D of Chapter 12, let me call attention to pp. 387–394, especially to item (3) on p. 392 and to the concluding statement concerning the $L-F$ contraction hypothesis on p. 394. The account of the testability of this hypothesis given there is superseded, as noted on p. 723 of Chapter 21, by the revised presentation in Section II of Chapter 21. In particular, the tempering in Chapter 21 of the claim of high disconfirmation made in Chapter 12 accords with the Duhemian cautions espoused in Chapter 17.

In his Panel article, M. G. Evans has offered criticisms of both Chapter 12 and 21 in regard to the treatment of the $L-F$ contraction hypothesis.⁷⁴ But, in my view, these have been effectively answered by H. Erlichson.⁷⁵

⁷⁰ A. Einstein, 'Relativitätsprinzip und die aus demselben gezogenen Folgerungen', *Jahrbuch der Radioaktivität* 4 (1907), pp. 412–413.

⁷¹ G. Holton, *op. cit.*, p. 180.

⁷² See Section 1.2, entitled 'Popper, Grünbaum and Holton on Lorentz and Einstein', p. 105, and also p. 99, of E. Zahar, 'Why did Einstein's Programme supersede Lorentz's? (I)', *British J. Philosophy of Science* 24 (1973), pp. 95–123.

⁷³ G. Holton, 'Mach, Einstein and the Search for Reality', in *Ernst Mach: Physicist and Philosopher*, [Boston Studies in the Philosophy of Science (ed. by R. S. Cohen and R. J. Seeger), Vol. VI], Reidel, Dordrecht, 1970, pp. 165–199.

⁷⁴ M. G. Evans, 'On the Falsity of the Fitzgerald-Lorentz Contraction Hypothesis', *Philosophy of Science* 36 (1969), pp. 354–362.

⁷⁵ H. Erlichson, 'The Lorentz-Fitzgerald Contraction Hypothesis and The Combined Rod Contraction-Clock Retardation Hypothesis', *Philosophy of Science* 38 (1971), pp.

§37

Finally, note that the magnitude of the metrogenic 'Einstein contraction' mentioned on p. 402 of Chapter 12 does, of course, correspond to the standard STR definition of simultaneity specified by the value $\epsilon = \frac{1}{2}$ in Equation (1) on p. 353. A generalization of that contraction as a function of ϵ ($0 < \epsilon < 1$) is derived in my book GCPP, Ch. III, §8.4, pp. 333–336. But the concluding sentence of the latter §8.4 regarding Bridgman's method of clock synchronism is to be understood in the light of the detailed philosophical analysis of that method in Chapter 20.

CHAPTER 14

§38

Page 420 contains an *incorrect* assertion of philosophic import concerning the Schwarzschild solution of Einstein's field equations: The claim, which lingers on in some treatises on the GTR,⁷⁶ that the derivation of the Schwarzschild solution depends on *separately requiring* that the desired radially symmetric solution be asymptotic to the Lorentz-Minkowski metric at large distances ('at infinity'). But as is pointed out in other treatises,⁷⁷ it is now more widely known from Birkhoff's theorem that a spherically symmetric GTR gravitational field in empty space *must* be asymptotic to the Lorentz-Minkowski metric 'at infinity', besides being static (time-independent). This point is overlooked by Glymour in his criticism of my views.⁷⁸

§39

To supplement the literature on Mach's Principle cited in Chapter 14, let me mention Goenner's study whose "purpose is to review, in the framework of the present-day understanding of Einstein's theory, the changes in interpretation which Mach's ideas have experienced and to sort out different statements collectively labeled as Mach's principle."⁷⁹ Also, there is an unpublished paper by Earman.⁸⁰

605–609. Additional material supporting my claims against Evans' objections can be found on p. 76 of L. Silberstein, *The Theory of Relativity*, Macmillan, London, 1914.

⁷⁶ See ch. 6, pp. 164–165 of R. Adler, M. Bazin and M. Schiffer, *Introduction to General Relativity*, McGraw Hill, New York, 1965.

⁷⁷ See S. Weinberg, *Gravitation and Cosmology*, John Wiley, New York, 1972, p. 337.

⁷⁸ C. Glymour, 'Physics By Convention', *Philosophy of Science* 39 (1972), p. 335.

⁷⁹ H. Goenner, 'Mach's Principle and Einstein's Theory of Gravitation', in R. S. Cohen and R. J. Seeger (eds.), *Ernst Mach: Physicist and Philosopher*, op. cit., pp. 200–215.

⁸⁰ J. Earman, 'Einstein and Mach: Covariance, Invariance, and the Equivalence of Frames; and the Special and General Principles of Relativity', (unpublished).

- §40 For statements which take issue with some of my views amid advocating the absolutistic as opposed to the relational conception of space-time and/or of its spatial and temporal submanifolds, see Hooker,⁸¹ Lacey,⁸² and Earman.⁸³ A reply to the pertinent part of Earman's paper is included in §3 of Chapter 22.
- §41 This completes the commentary on the chapters of the first edition. As shown in the Table of Contents, the next seven Chapters 16–22 are arranged in the topical order of the three Parts I, II and III upon which they respectively elaborate.

CHAPTER 16

- §42 First published in 1970, this chapter is the first installment of my response to 'A Panel Discussion of Grünbaum's Philosophy of Science'⁴ and supplements the first three chapters of Part I of the first edition. In view of the monograph-length of Chapter 16, the Abstract and Introduction on pp. 449–456 are followed on p. 457 by a fairly detailed table of contents of the remainder of the chapter.

I wish to call attention to three clarifications and/or corrections as follows:

- §43 (1) On p. 546, in the remarks on the role of geodetic *test* particles in the GTR, the sentence "A free particle of non-zero rest mass and also a photon (which has equivalent mass) has a geodesic path in *Minkowskian* spacetime" should be expanded to read: "... in *Minkowskian* space-time when seen as a particular solution of Einstein's GTR VACUUM field equations." One of the ramifications of this clarification is the following.

The STR is the theory of Minkowski or flat space-time, as noted in the above §5 Commentary on Sect. B of Chapter 2. Also, the GTR asserts that the space-time trajectory of a gravitationally monopole free *test* particle and of a free photon is a *geodesic* in either curved or flat space-time. And these geodetic test particles play a fundamental role – as explained in detail in Chapter 22, §2, Section (iii) – in the ontology of the

⁸¹ C. A. Hooker, 'The Relational Doctrines of Space and Time', *British J. Philosophy of Science* 22 (1971), pp. 97–130.

⁸² H. M. Lacey, 'The Scientific Intelligibility of Absolute Space', *British J. Philosophy of Science* 21 (1970), pp. 317–342.

⁸³ J. Earman, 'Who's Afraid of Absolute Space?', *Australasian J. Philosophy* 48 (1970), pp. 287–319.

GTR metric as formulated via Weyl's geodesic method. Hence the STR can be construed as a subtheory of the GTR in the sense that Minkowski space-time is a particular solution of Einstein's *vacuum* field equations to which the geodesicity assertions of the GTR concerning free test particle paths in space-time apply. Combine *this* understanding of the GTR and STR status of Minkowski space-time with the role of free *test particles* in specifying the metric of this FLAT space-time via Weyl's GTR geodesic method as detailed in Chapter 22, §2 (iii). Then the latter combination is seen to undercut and, I believe, to dispose of Glymour's criticism⁸⁴ of my claim on p. 546 regarding the logical overlap between the GTR and the STR with respect to the geodesic behavior of free test particles in Minkowski space-time.

Incidentally, the article by Earman cited under reference [16] on p. 546 and listed as 'forthcoming' in the bibliography has since been published.⁸⁵

§44

(2) On p. 547, in the parenthetical sentence in lines 4-6, the phrase "even when the energy-momentum tensor does *not* vanish" should be amplified to read "even when, in the standard (classical) *pre*-Wheeler GTR treatment of 1915, the energy-momentum tensor does not vanish." On the same page, in lines 19 and 20, the phrase "quite apart from any" should read instead: "even before we confront the."

§45

(3) On p. 563, my concession that "the GTR effects a normatively unique partition of its space-time intervals for a given matter-energy distribution" needlessly goes too far, as illustrated by the following example: J. Ehlers and W. Kundt have pointed out that besides Minkowski space-time there are an infinitude of known solutions of Einstein's *vacuum* field equations which do *not* admit the same metrical partition of space-time intervals into equivalence classes.⁸⁶ And they show that "there exist complete solutions of Einstein's vacuum field equations; that is, complete solutions free of sources (singularities), proving that it is permissible to think of a graviton field independent of any matter by which it be generated. This corresponds to the existence of source-free photon fields in electrodynamics."⁸⁷

⁸⁴ C. Glymour, 'Physics By Convention', *op. cit.*, p. 336.

⁸⁵ *General Relativity and Gravitation* 1, No. 2 (1970), pp. 143-157. For my brief reaction to this article, see my note by the same title in Vol. 2, No. 3 (1971), pp. 281-284.

⁸⁶ J. Ehlers and W. Kundt, 'Exact Solutions of the Gravitational Field Equations', in L. Witten (ed.), *Gravitation*, John Wiley, New York, 1962, p. 61 and Section 2.4, pp. 82ff.

⁸⁷ *Ibid.*, p. 97.

§46 The tantalizingly cryptic remarks on pp. 467 and 503–504 concerning Clifford's 'space-theory of matter' and Wheeler's geometrodynamical elaboration of it via the GTR have elicited criticism from Glymour, as has Chapter 16 generally.⁸⁸ Hence the much more detailed philosophical discussion of the Clifford-Wheeler theory given in Chapter 22, §3 includes a full rebuttal (§3 (b)) to Glymour's objections to pp. 503–504 of Chapter 16, which pertain to the ontology of curvature properties in the Clifford-Wheeler theory. And §3 (a) of Chapter 22 (p. 755) includes my reaction to Glymour's strictures regarding the doubts I had raised on p. 467 of Chapter 16 with respect to the capability of the Clifford-Wheeler theory to individuate its world-points.

§47 A number of queries and criticisms that have been presented in reviews of my GCPP have been answered indirectly, I believe, as part of the constructive, non-polemical development of ideas in Chapter 16.⁸⁹ Let me give one such example, because the reviewer remains unconvinced on the point in question. It pertains to the *revised* articulation in Chapter 16 of the GCPP treatment of the Riemannian idea which Weyl had put as follows:

I must begin by remarking that Riemann contrasts *discrete* manifolds, i.e., those composed of single isolated elements, with *continuous* manifolds. The measure of every part of such a discrete manifold is determined by the *number* of elements belonging to it. Hence, as Riemann expresses it, a discrete manifold has the principle of its metrical relations in itself, *a priori*, as a consequence of the concept of number.⁹⁰

The review in question is a sympathetic review-article by Arthur Fine, written before the first (1970) appearance in print of Chapter 16 but

⁸⁸ C. Glymour, 'Physics By Convention', *op. cit.*, pp. 322, and 338–339.

⁸⁹ I have in mind such reviews of GCPP as the following: E. W. Adams's in *Philosophy of Science* 39 (1972), pp. 553–555, to which Ch. 20 is also relevant; J. Merleau-Ponty's in *Rev. Philosophique* 160 (1970), pp. 471–475; and L. Sklar's in *Phil. Rev.* 81 (1972), pp. 506–508, as well as Sklar's 'The Conventionality of Geometry', *Am. Phil. Quart.*, monograph III (1969), pp. 42–60. Michael Friedman's 'Grünbaum on the Conventionality of Geometry' in *Synthese* 24 (1972), pp. 219–235 is directed, he says (*ibid.*, p. 235, n. 1), only toward my GCPP of 1968. But Friedman expresses his belief (*idem.*) that his paper's main criticisms apply to my later position in Ch. 16 of this book as well. The analysis he gives of GCPP is confined to its Ch. I. (As noted in the Preface of GCPP, this chapter dates from 1962 and was reprinted in GCPP for the convenience of the reader of Ch. III, which is my Reply to Putnam's critique of Ch. I).

⁹⁰ H. Weyl, *Space-Time-Matter*, Dutton, New York, 1922 (reprinted in 1950 by Dover, New York), p. 97.

published thereafter (1971).⁹¹ In the following quotation from that review-article, Fine therefore speaks of the 1970 essay reprinted in Chapter 16 as a 'forthcoming paper', saying:

... but we can also ask how cardinality could be relevant to an intrinsic metric. Suppose that we have a metric for a discrete space based on cardinality. In the simplest case we should assign as the length of an interval one less than the number of points in the interval. To judge the number of points, however, is to judge that a one-to-one correspondence can be established between the points of the spatial interval and some appropriate counting class. We do this usually by counting or by some other physical matching process. The point I want to make, however, is not about how we determine cardinality, but it is about the concept itself. The very idea of cardinality involves the idea of a one-to-one correspondence between two classes. Even if both range and domain of the correspondence are sets of spatial points (and why not?), the correspondence itself is external, both conceptually and in practice, to the spatial framework. Thus if I must judge on the basis of cardinality, whether two disjoint spatial intervals are congruent I must imagine some process or procedure outside of the intervals themselves by virtue of which I can attempt to make a one-to-one correspondence between the given intervals. There are a variety of ways to try, but they are all external. It is no more a 'built-in' feature of an interval that it contains ten space atoms, than it is a built-in feature that it admits of ten applications of my standard rod. Conceptually, I am in both cases dealing with a relation between a given spatial interval *and something else*. For finite cases one can, to be sure, attempt to construe cardinality in the manner of Russell, that is, one can express statements of cardinality by means of purely logical formulas that involve only quantification and identity. It is not clear, however, just what bearing this sort of reduction has on my claim that the concept of cardinality (as well as that of equi-cardinality) is conceptually linked with the concept of a one-to-one correspondence. (In the forthcoming paper cited in my leading note, Grünbaum argues that the dyadic property of equicardinality is intrinsic. I intend my claim here to challenge that argument.) If this claim is correct then insofar as Grünbaum constrains the admissible metrics so that for discrete space the metric must depend on the cardinality of intervals, then he will have lost the connection he wanted between *intrinsic* and *built-in* as well as the contrast between *intrinsic* and *external*.⁹²

This criticism embodies Fine's reaction to a letter of mine concerning only the first draft of his objection. Let me therefore now divide my comment on the criticism as published into two parts as follows: (A) First, I shall state the substance of the response I offered to Fine's first draft, and (B) I shall amplify this response by dealing with the argument Fine gives for having found my initial reply unsatisfactory.

(A) In the Ch. 16 treatment, I consider metrics — be they non-trivial or trivial in the sense of its §2 (b) — which generate the same partition

⁹¹ A. Fine, 'Reflections on a Relational Theory of Space', *Synthese* 22 (1971), pp. 448–481; to be reprinted in a forthcoming volume *Space, Time and Geometry*, edited by P. Suppes and published by Reidel.

⁹² *Ibid.*, p. 460.

of the class of intervals into equivalence classes as does the dyadic property of equi-cardinality possessed by intervals. And I deemed such metrics intrinsic. The issue posed by Fine is: Does the concept of one-to-one correspondence make for externality, thereby rendering the equi-cardinality of two intervals a dyadic property *external* and hence extrinsic to the *pair* of intervals related by it? Being mindful of the definitions on pp. 505–506, especially of the definition (8) of the ‘externality’ of a dyadic property, I cannot see that the avowed logical involvement of the notion of one-to-one correspondence in the concept of equi-cardinality is such as to render the *dyadic* property of equi-cardinality external to the *PAIR* of intervals possessing it. For the Chapter 16 account rests the intrinsicality of a cardinality-based metric on the non-externality (and ‘generality’ à la (3) on p. 505) of the following property: the equi-cardinality sustained *TO ONE ANOTHER* by a *PAIR* of intervals, *NOT* on some relation they bear to something external to both of them!

Though superseded by the Chapter 16 account, even the GCPP treatment would seem to be immune to Fine’s particular criticism here. Let us see why I think so.

The GCPP presentation rested on treating cardinality as an intrinsic *MONADIC* property of intervals, as distinct from the Chapter 16 reliance on equi-cardinality as an intrinsic dyadic property. Fine’s point regarding the monadic approach in GCPP is, I take it, that the cardinal number of a set is the set of all sets equi-potent (equi-pollent) with the given set, so that the concept of cardinal number involves one-to-one correspondence with *OTHER* sets *EXTERNAL* to the given set. But, to my mind, this critique overlooks that whereas the *generic* concept of cardinal number does involve such one-to-one correspondence, metrics which make the measure of an interval a function of its cardinality can nonetheless be specified in such a way as to *avoid* precisely such an *external* kind of one-to-one correspondence. Let us consider the measures $M(S)$ of sets S which are a function of their cardinality $\bar{S}$ for finite and transfinite cardinalities in turn.

(1) $\bar{S} = n$, where n is a finite integer. To confer specificity on the case mentioned in Weyl’s *aforecited* statement, consider a physical space which is discrete in the more restricted sense of Chapter 16, p. 462.⁹³ And

⁹³ Cardinality-based metrics on such a space are discussed, for example, on pp. 485 and 490 of Ch. 16.

let ' $nc(S)$ ' be short for 'the cardinal number of the set S '

Then I am able to say — via a Russellian, purely logical translation of the statement that $nc(S) = n$ — that S has n members WITHOUT any recourse to saying that S is mappable one-to-one onto certain OTHER sets. It is certainly ALSO true that S is thus mappable. But contrary to Fine, such conceptual linkage with external mappability as obtains on the strength of this fact does *not* make such external mappability a logically indispensable presupposition of the statement $nc(S) = n$, so as to render it a NON-intrinsic monadic property of the set S to have n members. Thus, in my erstwhile monadic approach to the specification of the metric for a discrete space, I do *not* make indispensable use of the cardinal number of set S in the sense of pre-supposing an appeal to one or all sets equi-potent with it. Hence even if my definitions (9) and (10) on p. 507 had not disavowed my former monadic approach of GCPP, I would *not* need to rely on properties *external* to the one given interval, in the sense of 'external' specified for monadic properties by my new definitions (2) and (1) on p. 505 of Chapter 16.

Still pursuing my erstwhile GCPP monadic approach for now, we next consider intervals having transfinite cardinalities.

(2) $\bar{S}$ is transfinite. Under this heading, we need to consider dense spaces whose non-degenerate intervals are either denumerably infinite or have the cardinality of the continuum. Let ' $\{a\}$ ' represent any singleton, and ' A ' any non-degenerate interval. Then, as noted on pp. 526 and 531–532 of Chapter 16, the *cardinality-based* measure $M(\{a\}) = 0$, $M(A) = +\infty$ is, of course, *trivial* in the sense of §2 (b) of Chapter 16, while being countably additive in continuous space and only finitely additive in dense denumerable space.

Am I justified in deeming this metric to be intrinsic in the face of Fine's charge of externality? Having given a Russellian logical translation of ' $nc(S) = 1$ ' in order to be able to characterize a singleton ' $\{a\}$ ' as such, I can now handle the respective assertions that a set A is denumerably infinite or has the cardinality of the continuum *without* any recourse to properties *external* to an interval possessing either transfinite cardinality. Thus, I invoke *only* such one-to-one correspondences as do or do not obtain between the given interval *and its OWN subsets* as follows: I can define ' A is an infinite set' à la Dedekind-Cantor by saying that there is a one-to-one mapping of A onto a proper subset of itself. And I can define

' A is denumerably infinite' by saying that every infinite subset of A can be mapped one-to-one onto A . And IF the cardinality of the continuum is the next higher one after aleph zero, I can define ' Y has the cardinality of the continuum' in the following manner for which I am indebted to Gerald Massey: $(\exists x) [x \text{ is a subset of } Y, \text{ and } x \text{ is infinite, and } x \text{ cannot be mapped one-to-one onto } Y, \text{ and } (z) (\text{if } z \text{ is an infinite subset of } Y, \text{ then either } z \text{ can be mapped one-to-one onto } x \text{ or } z \text{ can be mapped one-to-one onto } Y)]$.

This brings me to my amplification of the preceding response to Fine's first draft. I now need to deal with the argument he gives in his published text for not having found my initial reply satisfactory.

(B) The nub of Fine's challenge to my Chapter 16 claim that the dyadic property of equi-cardinality is intrinsic and hence non-external is his contention that the presupposed one-to-one correspondence is external for the following aforesaid reason: "Even if both range and domain of the correspondence are sets of spatial points (and why not?), the correspondence itself is external, both conceptually and in practice, to the spatial framework."

Here Fine contends that the equi-cardinality of two space intervals fails to be intrinsic not only to this pair of intervals but even to the 'spatial framework' of which they are a part (cf. p. 509 of Chapter 16 for my definition (14) of the *latter* kind of intrinsicity). For the sake of specificity, let us consider pairs of intervals which are either disjoint or 'quasi-disjoint' in the sense of Chapter 16, p. 528. Then I do not understand Fine's claim of extrinsicity to the spatial manifold as a whole: We can think of the one-to-one correspondence as being effected *within* the spatial framework by connecting lines which pair off the points of the two intervals with one another in any one of an infinitude of alternative ways, no matter which. How then does Fine reason that the correspondence *itself* is external, both conceptually and in practice, to the spatial framework as a whole? Furthermore, if the union of two disjoint intervals were itself an interval constituting the whole of the space, would Fine say that the requisite one-to-one correspondence effected by the connecting lines is *external* to the spatial framework or even to the pair of intervals?

If equi-cardinality is to be held a dyadic property *external* to the pair of intervals possessing it on the ground that the presupposed one-to-one

correspondence is itself thus external, then I ask: How would Fine construe the presumed externality of a *dyadic* property so that the *mere dyadicity* of any dyadic property P_2 of a pair of disjoint or quasi-disjoint intervals does not AUTOMATICALLY entail P_2 's externality? In other words, Fine may be tacitly claiming that the class of *dyadic* properties intrinsic to a pair of such intervals is simply *empty* in virtue of their dyadicity. If so, how can any construal of externality non-vacuously allow for the very contrast between intrinsicity and externality which, he tells us, was lost by my claim that *equi-cardinality* is an intrinsic dyadic property?

If Fine were to reply that this is my problem and not his, I can say only the following: I do not see how my denial of this loss of contrast can be gainsaid by Fine's proposed assimilation of the one-to-one correspondence that makes for equi-cardinality to the *externality* of the property possessed by two intervals in virtue of their permitting the same number of external applications of a standard rod. It was the belief that equi-cardinality does sustain the contrast claimed by me which constitutes one of the key ideas in Riemann's Inaugural Dissertation, as stressed in the above quotation from Weyl. But it is to be borne in mind, as I emphasize on pp. 526-527 of Chapter 16, that neither Riemann nor Weyl were interested in cardinality-based metrics defined on *dense* spaces precisely because these metrics were trivial even though they are intrinsic.

CHAPTER 17

§48 As noted in the Commentary on Chapter 4, in Chapter 17 my erstwhile use of physical geometry as a crucible in which to examine Pierre Duhem's influential epistemological holism for physics is substantially broadened and revised partly in response to criticisms that were made of the Chapter 4 treatment.

§49 Perhaps some remarks on my closely related forthcoming book *Falsifiability and Rationality* (University of Pittsburgh Press) are here in order.

In that book, I try first to provide a genuinely fresh critical examination of Karl Popper's conception that the falsifiability requirement is the hallmark of scientific rationality to the exclusion of inductive confirmability. Popper's account of post-Baconian inductivism is challenged

by evaluating the scientific status of Freudian psychoanalysis according to both inductivist and Popperian canons: An analysis of the current status of the scientific credentials of psychoanalytic psychotherapy is claimed to show that the latter can be impugned by inductivist criteria no less than by Popper's falsificationist criterion. Thus, I argue there that inductivism has the logical resources for demanding controlled studies of therapeutic efficacy and can demonstrate why the remission of the neurotic symptoms of psychoanalytically treated patients need not at all redound to the credit of Freudian therapeutic theory. I also address myself to other theses of Popper's which turn on his claim that falsifiability should be accorded a distinguished epistemic role in science as contrasted with confirmability (corroborability).

Though not sharing Popper's version of falsificationism, I attempt to offer an alternative constructive conception of falsifiability which refines some formulations of Chapter 17. In particular, I present an account of the logical conditions under which strong presumptions of falsity for scientific hypotheses are rationally warranted. In doing so, I am cognizant of the challenge posed not only by Duhem's writings but also by some highly counterintuitive results of Rudolf Carnap's inductive logic as elaborated by Wesley Salmon. And I thereby continue to reject Imre Lakatos's universal falsificationist agnosticism in regard to the following contention of his: Empirical results cannot make it rationally mandatory to presume that a given scientific hypothesis is false, any more than different test outcomes could instead have required the presumption that it is true. But it is there seen to emerge from the corpus of Lakatos's writings that his position is not vulnerable to the rhetorical question I asked in n. 64 on p. 618 of Chapter 17.

There is also a discussion of the bearing of the rational credentials of scientific hypotheses on scientific research strategies or serendipity and on understanding the history of science.

CHAPTER 18

- 50 This chapter dates from 1969 and goes beyond both editions of my MSZP. Its focus for dealing with Zeno's paradoxes of *motion* anew is Hermann Weyl's contention that if a machine obeying the principles of classical kinematics cannot carry out a denumerable infinity of

distinct operations in a finite time, then the received interpretation of the classical mathematical theory of motion is beset by one of Zeno's kinematical paradoxes.

In the context of classical kinematics, allegations of paradox against the traversability of a finite space interval in a finite time have been leveled in the literature inspired by Zeno against at least the following two kinds of motions:

(I) One uninterrupted motion which *can be analyzed* into an infinite number of successive submotions, as in the case of the *legato* run of Zeno's Achilles discussed on pp. 632–637 of Chapter 18, and

(II) An intermittent motion consisting of N_0 motions of suitably decreasing durations, separated by N_0 suitably decreasing pauses of rest, as in the case of the motion of my *staccato* runner of pp. 638ff.

Let me refer to these two kinds of motions as 'type I' and 'type II' respectively. Then I can say that Chapter 18 endeavors to carry forward the arguments of MSZP by buttressing the following contentions further:

(i) *Some* designs for motions of type I no less than *some* designs for motions of type II are indeed vitiated by kinematic paradoxes in the sense of requiring *discontinuities* (either finite or infinite) in at least one of the functions of the time variable which respectively specify the moving body's position, velocity and acceleration, and (ii) For some motions of type II no less than for some of type I, there are *other* designs which demonstrably do *not* involve any such discontinuities, and (iii)

Allegations of paradox that have been made against these *latter* designs are gratuitous and ill-founded but derive their plausibility from the tacit appeal to the positive threshold or minimum which limits direct, unaided human awareness of temporal durations. Commenting on these contentions, as presented in my contribution to Wesley Salmon's aforecited 1970 anthology *Zeno's Paradoxes*, the physicist Philip Morrison expressed the following view:

His [Grünbaum's] methods go more deeply into the mathematical theory of the continuum than those of his predecessors.... By designing his imaginary mechanisms so that no infinite distances need to be traveled by any lever or cog, he demonstrates that supertasks, like the task of Achilles, can indeed be accomplished, not only in the context of the steady race but also very widely. No contradictions, logical or kinematic, appear at any point in Grünbaum's piece. That is not to say, of course, that such machines might be built to order: they remain in an abstract nonatomic world, a world satisfying which our world

can do only by postulate so far the full demands of measure theory. That is all one can ask for the present.”⁹⁴

On the other hand, the mathematician J. Q. Adams’s 1973 article ‘Grünbaum’s Solution to Zeno’s Paradoxes’,⁹⁵ is intended to offer a counterexample to my MSZP, to my paper ‘Are ‘Infinity Machines’ Paradoxical?’,⁹⁶ and to Chapter 18. My reply to Adams’s article appears in the same issue of *Philosophia*, pp. 51–57.

CHAPTER 19

- §51 As pointed out in n. 1 of this chapter, Eugene Wigner is reported to have declared that ‘Entropy is an anthropomorphic concept’. Since the statistical entropy of a physical system played a role in the account of time’s arrow in Chapter 8, Chapter 19 addresses itself to the appraisal of Eugene Wigner’s dictum in the context of the coarse-grained entropy of classical statistical mechanics. One aim of Chapter 19 is to provide an ontological underpinning for Section A, subsection II of Chapter 8 in the sense of arguing that the coarse-grained entropy of classical statistical mechanics can be construed in scientific realist fashion instead of being a mere anthropomorphism, as has been charged.

CHAPTER 20

- §52 Pp. 670–708 of this chapter appeared originally as an article of a 1969 Panel Discussion of simultaneity by infinitely slow clock transport in the STR and GTR. Wesley Salmon and I jointly wrote the brief Introduction to this Panel Discussion. And since the latter sets the stage for my own Panel essay, Chapter 20 begins with that Introduction. The Panel Discussion was largely prompted by Ellis and Bowman’s philosophical gloss on P. W. Bridgman’s 1962 method of synchronism by infinitely slow clock transport as an alternative to Einstein’s light signal method in the STR. Since the four Panel papers addressed themselves in part to a critique of the Ellis and Bowman paper, the reader is referred to the

⁹⁴ P. Morrison, *Scientific American* (March 1971), p. 123.

⁹⁵ *Philosophia* (Israel) 3 (Jan. 1973), pp. 43–50.

⁹⁶ *Science* 159 (1968), pp. 396–406.

spirited reply by Ellis.⁹⁷ It should be noted that Ellis's reply was written before November 1970, and hence before the publication of Chapter 16 in December 1970.

§53 The formulations in Chapter 16 supplant the 1969 statements on Riemann-conventionality made on pp. 690–691 and thereafter in Chapter 20.

Several clarifications and/or emendations in the text of Chapter 20 are in order.

§54 (1) Having formulated the rule *RCS* on p. 672, I stated a certain 'generalized form' of it on p. 673 and denoted the latter form by '*GRCs*'. Howard Stein has offered the following elucidating comment on the logical relation between *RCS* and its generalization *GRCs*.

RCS disallows 'correcting' the reading of any clock whose readings have already been used in synchronizing another clock. *GRCs*, in effect, disallows 'correcting' the reading of any clock that has once been synchronized. For it requires 'at most one setting per clock', and 'correcting' is tantamount to resetting. If, in applying this rule, we count the arbitrary initial setting of the first clock as a 'synchronization' as intended, and if we make the reasonable stipulation that no 'unsynchronized' clock will be used in synchronizing another clock, then *RCS* follows from, but does not entail, *GRCs*.

§55 (2) As given on pp. 674–676, the proof of the intransitivity of non-standard synchronisms does not explicitly cover the special case of three clocks stationed at as many *collinear* points. The deduction on p. 675 assumes *non-collinearity* by employing the premise $l + m - n \neq 0$. But Philip Quinn has shown that my argument can also be used to cover the case of three collinear clocks provided that the optical synchronization procedure which was implicit in that argument all along is used.⁹⁸ Quinn calls this optical procedure 'the two-ray method'. And Quinn points out that there is an alternative optical procedure for synchronizing clocks – the 'one-ray method' – with respect to which all non-standard synchronisms are required to be transitive. The moral of Quinn's paper is that (i) there are at least two procedures for synchronizing triplets of clocks in an inertial system, (ii) these alternative procedures impose

⁹⁷ B. Ellis, 'On Conventionality and Simultaneity – A Reply', *Australasian J. Philosophy* 49 (1971), pp. 177–203. The last page contains the entry 'Received November 1970'.

⁹⁸ P. L. Quinn, 'The Transitivity of Non-Standard Synchronisms', unpublished.

different constraints on the transitivity properties of *non*-standard synchronisms, so that (iii) assertions about the transitivity properties of these synchronisms need to be relativized to the synchronization procedure(s) for which they hold.

§56 (3) The defense I offer on pp. 677–678 of the admissibility of the particular directionally-dependent synchronization rule ε_ϕ proposed by Reichenbach⁹⁹ is intended solely illustratively and affects nothing of substance in that chapter. Nonetheless, I should mention an inadvertence there which has been pointed out by Clark Glymour.

Recall from item (1) in §54 above Howard Stein's explanation of how *GRCS* entails *RCS* but not conversely. Thus, to reject *RCS* is to disavow *GRCS* as well. But the disavowal of the logically stronger *GRCS* does not, of course, entail the denial of the weaker *RCS*. Since Ellis and Bowman impose not only the requirement *RCS* but also the stronger *GRCS*, they mean to use the fulfillment of both of these requirements as a touchstone for assessing the acceptability of the non-standard kind of synchronism ε_ϕ proposed by Reichenbach. And, in effect, they indicted his ε_ϕ for violating *GRCS*. Now, on pp. 677–678, I endeavored to exonerate Reichenbach in the sense that Ellis and Bowman's case against him altogether begs the question.

But to make such an exoneration cogent, it is not enough to point out, as I do on p. 677, that Reichenbach disavowed the stronger *GRCS*, and hence cannot be indicted – without begging the question – for having proposed his ε_ϕ in violation of *GRCS*. Cogency of rebuttal against Ellis and Bowman demands that if Reichenbach's ε_ϕ violates the weaker *RCS* as well, then I must show also that he rejected *RCS* no less than *GRCS*. Glymour observed, in effect, that by violating Equation (5) on p. 672, which is deduced from *RCS*, Reichenbach's proposed ε_ϕ does violate not only *GRCS* but also *RCS*. Hence, in order to exonerate Reichenbach, it was incumbent upon me to adduce as well Reichenbach's rejection of *RCS*, which I failed to do. However, it seems rather clear that Reichenbach would not have agreed to *RCS* as a generally valid requirement, any more than he did to *GRCS*.

§57 I should explain why I referred on p. 673 to Ellis and Bowman's tacit

⁹⁹ H. Reichenbach, *The Philosophy of Space and Time*, Dover, New York, 1958, p. 162, Equation (2). On p. 163, Reichenbach calls this formula 'an admissible definition of simultaneity'.

use of *RCS*, as distinct from *GRCS*, by calling it 'a matter of small importance'. This valuation was intended to be confined to the role played by *RCS* on p. 672 in Ellis and Bowman's enthymematic deduction of Equation (5) from the two-way light principle

§ 58 (4) On p. 693, the following paragraph should be inserted, so as to precede immediately the last paragraph on the page, which begins with "Referring back to our world-line diagram . . .": Like Ellis and Bowman, I shall assume that each of the moving clocks *U* itself qualifies as an inertial system no less than the system *I* whose clocks are to be synchronized or that each clock *U* is a *free* 'particle' in *I*. My impending claims concerning the readings of the clocks *U* are predicated on this assumption! Bridgman's original method ([2], p. 65) assumed that the '*self-measured*' velocities of the various transported clocks were each *constant*.

§ 59 A number of issues discussed in both Chapter 20 and Chapter 12 are treated illuminatingly by Winnie.¹⁰⁰

§ 60 Finally, in Chapter 20 I did not discuss the so-called STR clock paradox because I regard it as having been quite adequately resolved in the literature.¹⁰¹

CHAPTER 21

§ 61 I already discussed this supplement to Chapter 12 in the Commentary on the latter vis-à-vis some of Holton's views. A June 1973 symposium on Space, Matter and Motion, held at the Ohio State University (Columbus) has produced additional relevant literature.¹⁰²

CHAPTER 22

§ 62 One of the major omissions of this chapter is the philosophic import of

¹⁰⁰ J. Winnie, 'Special Relativity without One-Way Velocity Assumptions', *Philosophy of Science* 37 (1970), Part I, pp. 81–99; Part II, pp. 223–238.

¹⁰¹ See, for example, Y. P. Terletskii, *Paradoxes in the Theory of Relativity*, Plenum Press, New York, 1968, Ch. III, §12, pp. 38–41; C. B. Lefvert and T. M. Donahue, 'Clock Paradox and the Physics of Discontinuous Gravitational Fields', *Am. J. Phys.* 26 (1958), pp. 515–523; and N. D. Mermin, *Space and Time in Special Relativity*, McGraw-Hill, New York, 1968, Ch. 16.

¹⁰² Cf. K. Schaffner, 'Space and Time in Lorentz, Poincaré, and Einstein: Divergent Approaches to the Discovery and Development of the STR' (forthcoming), and W. C. Salmon, 'Clocks and Simultaneity in Special Relativity', (forthcoming).

recent work on the inevitability of gravitational collapse and on the necessary incompleteness of the space-time manifold for a wide class of solutions of the Einstein equations.¹⁰³ Concerning these developments, Dieter Brill has expressed the judgment that "It would be difficult to find a topic of more impact on philosophical ideas about space and time than these singularity theorems."¹⁰⁴

§63 On the other hand, §§2 and 3 of Chapter 22 include a scrutiny of the conception of 'the real geometry of space-time' espoused in John Graves's 1971 book *The Conceptual Foundations of Contemporary Relativity Theory*. By providing an examination of this conception of Graves, some of the discussion in the chapter deals with the basis on which he rests his appraisal of my 'Reply to Putnam', given as part of a recent review by him.¹⁰⁵

§64 The theorem asserting CPT-invariance, which is adduced on p. 797 as the foundation for inferring *T*-invariance violation from 'observed' CP-invariance violation, has been stated by Weinberg as follows: "two observers, who distinguish particles and antiparticles oppositely *and* distinguish right and left oppositely, *and* distinguish the past from the future oppositely, will measure the same probabilities for corresponding events."¹⁰⁶ Technical details on the CPT conservation theorem can be found, for example, in books by Streater and Wightman,¹⁰⁷ and by Bernstein.¹⁰⁸

The *T*-invariance violation by the 2π -decay mode of long-lived neutral *K*-mesons which has been asserted via the CPT-theorem, is questioned

¹⁰³ Cf. S. W. Hawking, 'The Occurrence of Singularities in Cosmology', *Proc. Roy. Soc. London A* **294** (1966), pp. 511–521; R. Geroch, 'Local Characterization of Singularities in General Relativity', *J. Math. Phys.* **9** (1968), pp. 450–465; C. W. Misner, 'Absolute Zero of Time', *Phys. Rev.* **186** (1969), pp. 1328–1333; R. Penrose and S. W. Hawking, 'The Singularities of Gravitational Collapse and Cosmology', *Proc. Roy. Soc. London A* **314** (1970), pp. 529ff.; S. W. Hawking and G. F. R. Ellis, *The Large-Scale Structure of Space-Time*, Cambridge University Press, Cambridge, 1973, Chs. 8–10; and J. Earman, 'Till the End of Time' (unpublished).

¹⁰⁴ D. R. Brill, 'Space-Time Philosophy', *Science* **175** (1972), p. 53.

¹⁰⁵ J. C. Graves, Review of Vol. V of the *Boston Studies in the Philosophy of Science*, in *British J. Philosophy of Science* **24** (1973), pp. 183–185.

¹⁰⁶ S. Weinberg, 'Where We Are Now', *Science* **180** (1973), p. 277.

¹⁰⁷ R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, And All That*, W. A. Benjamin, New York, 1964, pp. 142–143.

¹⁰⁸ J. Bernstein, *Elementary Particles and Their Currents*, W. H. Freeman, San Francisco, 1968, Ch. 5.

by Mehlberg who writes: "The compatibility of time-reversal invariance with the behavior of electrically neutral K -mesons can also be shown, and is dealt with in my forthcoming work on *The Problem of Time*."¹⁰⁹ This book is not yet available at this writing. But I cannot see that Mehlberg's outline¹¹⁰ of his forthcoming detailed argument is sufficiently clear to gainsay the interpretation of the CP-violation experiments given in the Sachs article cited on pp. 797 and 802.

The inconclusiveness of the current evidence from particle physics for T -invariance *violation* is indicated by Dass's contention that "In certain respects, CPT-invariance has been verified experimentally to a poorer accuracy than CP-invariance and T -invariance."¹¹¹ In this sense, it would be imprudent to rest any part of the case for T -invariance *violation* on CPT-invariance.

¹⁰⁹ H. Mehlberg, 'Philosophical Aspects of Physical Time', *op. cit.*, p. 368. The title of Mehlberg's forthcoming book has been changed to the one cited in §29 of this Appendix under the commentary on Chapter 8.

¹¹⁰ *Ibid.*, pp. 368-370.

¹¹¹ Cf. G. V. Dass, 'CP, T and CPT Invariance in Neutral Kaon Decays', *Fortschritte der Physik* **20** (1972), p. 77. For the broader context of these issues, see D. I. Blokhintsev, *Space and Time in the Microworld*, D. Reidel, Dordrecht and Boston, 1973.

INDEX OF PERSONAL NAMES

(Compiled by Theodore C. Falk)

- Abraham, M., 374
 Achinstein, P., 594, 597, 599, 601, 602
 Adams, E. W., 843n.
 Adams, J. C., 575
 Adams, J. Q., 851
 Adams, N. L., 411n., 416n., 439
 Adler, R., 537, 566, 770, 772, 793, 801, 840n.
 Aharoni, J., 394n., 395n., 429, 604n., 727n.
 Akhiezer, N. I., 13n., 429
 Alexander, H. G., 418n., 429
 Alexander, H. L., 113n., 429
 Alfred the Great, 51, 55
 Anderson, A. R., xxiii, 607
 Anderson, J. L., 536, 540n., 566
 Aristotle, xiii, 158, 333, 382, 406, 429, 571, 835
 Arzeliès, H., 455, 566
 Augustynek, Z., 820

 Baker, B., 333n., 429
 Baldus, R., 157n., 429
 Barankin, E. W., 85n., 429
 Bar-Hillel, Y., 337n., 434
 Barker, S. F., 292, 298
 Baruch, J. J., 55n., 429
 Baum, W. A., 119n., 429
 Bazin, M., 537, 566, 770, 772, 793, 801, 840n.
 Beadle, G. W., 571, 574
 Bell, E. T., 379, 429, 664n.
 Belnap, N., xxiii, 450n., 607, 822
 Beltrami, E., 151, 784
 Berger, G., 823, 832
 Bergman, T. O., 596
 Bergmann, G., 244n., 429
 Bergmann, H., 184n., 288n., 322, 325, 429
 Bergmann, P. G., 360n., 371n., 394n., 429, 504, 566, 599n., 603n., 604n., 682n., 707, 772, 801
 Bergson, H., 251, 326, 429
 Bergstrand, E., 386n., 429
 Berkeley, G., 422, 430
 Bernays, P., 150, 153n., 436
 Bernal, J. D., 574n.
 Bernoulli, J., 166–167
 Bernstein, J., 855
 Bilaniuk, O. M. P., 822n., 827
 Birkhoff, G. D., 369n., 430, 662
 Black, M., 217, 326–329, 430, 631
 Blank, A. A., 154–157, 430, 435
 Blatt, J. M., 245n., 430
 Blokhintsev, D. I., 856n.
 Bludman, S. A., 687
 Blum, H. F., 251n., 430
 Boeder, P., 157n., 435
 Bohm, D., 250n., 430
 Bohr, N., 252–253
 Boltzmann, L., 229n., 239–240, 244, 246–247, 260, 265, 278–280, 790, 830
 Bolza, O., 13n., 430
 Bolzano, B., 159, 430, 484
 Bonch-Bruевич, A. M., 395n., 430
 Bondi, H., 221n., 321, 406–407, 430, 537, 566
 Bonola, R., 39n., 47n., 121n., 430
 Born, M., 244–248, 405n., 430
 Bourbaki, N., 484, 566
 Bowman, P., 668–673, 676–684, 688, 691–693, 696, 699–707, 851–854
 Boyer, C. B., 159n., 430
 Braithwaite, R. B., 122–123, 430
 Braude, A. I., 113n.
 Bridgman, P. W., 41–44, 134, 147, 159–160, 221–233, 260, 289, 354n., 368n., 431, 532, 547, 566, 667–669, 693, 695, 707, 810, 812, 816, 818, 840, 851, 854
 Brill, D. R., 855
 Brillouin, L., 253, 288n., 431
 Broad, C. D., 126
 Bromberger, S., 307
 Brouwer, C., 55n., 431

- Brown, F. A., Jr., 58, 59n, 431
 Brans, C. H., 422, 423, 430, 431
 Bruce, V. G., 58n, 439
 Büchel, W., 833
 Burington, R. S., 332n, 431
 Burley, W., 561

 Callaway, J., 421n, 431
 Camp, J., 589, 656, 664
 Campbell, N. R., 134
 Cantor, G., 10, 159, 176, 431, 484, 566, 818, 846
 Capek, M., 328, 431
 Caratheodory, C., 219, 220n, 431
 Carnap, R., 81, 105, 123n, 127, 134, 152, 153, 179, 278, 280, 431, 506, 783, 799, 820, 849
 Carter, B., 794, 801
 Cassirer, E., 319n, 431
 Cavalieri, B., 158
 Cech, E., 460n., 566
 Cedarholm, J. P., 398n, 432
 Chandrasekhar, S., 201, 202, 432, 824
 Cherry, E. C., 288n, 432
 Chihara, C. S., 631, 642, 643, 645
 Christoffel, E. B., 784
 Clarke, S., 418
 Clemence, G. M., 67n., 75n., 400n, 432
 Clifford, W. K., xxii, 41, 42, 424n, 432, 452, 467, 502, 504, 547, 566, 728, 730, 750, 754, 764, 765, 767, 783, 788, 801, 843
 Coffa, J. A., xxiii, 450n., 462, 832
 Cohen, R. S., xv, xxiii
 Copson, E. T., 333n., 429
 Costa de Beauregard, O., xviii, 227, 232, 248n., 266, 432
 Courant, R., 146n., 164n., 167n., 432
 Couturat, L., 128
 Cramér, H., 164, 166n., 172n., 432
 Creath, R., 664
 Crombie, A. C., 561, 566
 Csonka, P. L., 825n.

 D'Abro, A., xv, 12n., 56n., 68, 69, 432, 501, 553, 554, 566, 664n.
 Dass, G. V., 856
 Davies, R. O., 251n., 432
 Dedekind, R., 175n., 432, 846
 Delboeuf, J., 330, 331

 Demopoulos, W., 449, 499, 500, 511, 523, 566
 de Sitter, W., 420
 Dicke, R. H., 422, 423, 430, 432, 454, 572, 578, 583, 586, 614, 629n
 Dingle, H., 378, 394, 405n, 407n, 432
 Dingler, H., 103, 104, 129, 432
 Dirac, P. A. M., 400, 454, 725, 727, 789
 Donahue, T. M., 854
 Driesch, H., 325, 432
 du Bois-Reymond, P., 159, 161, 167, 432, 810, 816
 Dugas, R., 129n, 440
 Duhem, P. (*see also* D-theists), 106, 115, 123, 126, 131, 149, 383, 433, 584, 587, 601, 602, 606, 614, 625, 807, 808, 848, 849
 Durney, B., 629n

 Earman, J., xxiii, 449, 500, 535, 542, 544, 546, 547, 553, 555, 560, 562, 564, 566, 605, 754, 755, 775, 788, 791, 799, 800, 801, 822n, 823, 828, 831, 834, 840, 842, 855n
 Eddington, A. S., 24, 41, 125, 208n, 210, 221, 230, 265n, 289, 319, 321, 433, 464, 471, 567, 790
 Edel, A., 158n, 433
 Ehlers, J., xxii, 734, 735, 736, 737, 740, 749, 756, 764, 792, 801, 828, 842
 Ehrenfest, P., xiii, 241, 333n., 433
 Ehrenfest, T., 220n, 240, 433
 Earstein, H., xxiii
 Einstein, A., xvii, 8, 261, 274, 275, 336, 433, 752, 839; general theory of relativity and philosophy of geometry, 9, 22, 66, 79, 93, 106, 151 (esp. 131, 147), 201, 408, 418, 428, 450, 503, 544, 546, 551, 563, 567, 577n., 611, 613, 730, 731, 740, 746, 756, 762, 773, 787, 801, 824, 840, 841, 842; special theory of relativity and simultaneity, 28, 31, 61, 192, 334, 341, 417, 425, 451, 561, 598, 600, 666, 671, 675, 678, 681, 683, 692, 693, 696, 702, 707, 709, 717, 723, 726, 727, 834, 836, 837, 838, 839, 851
 Eisenhart, L. P., 13n., 102n., 120n., 139n., 150n., 433, 472, 474, 501, 567, 761, 768n., 801

- Ellis, B., 668, 673, 676, 684, 688, 691, 693, 696, 699, 707, 851, 854
 Ellis, G. F. R., 828n., 855n.
 Epstein, P., 241n., 405n., 434
 Erlichson, H., 839
 Essen, I., 400, 434
 Euclid, 158, 434, 744
 Evans, M. G., 839

 Faris, J. J., 833
 Feigl, H., xv, 77, 205
 Feinberg, G., 453n., 567, 687, 822n.
 Feyerabend, P. K., xv, 33, 34, 94, 292n., 336, 337, 434
 Feynman, R. P., 207, 253n.
 Fine, A., 808, 843, 848
 Fitzgerald, G. F., 387, 403, 408, 718, 725
 Fitzgerald, P., 825
 Fizeau, A. H. L., 355n., 374
 Fletcher, J. G., 424n., 434
 Foucault, J. B. L., 585
 Fox, R., 822n.
 Fraenkel, A., 175n., 337, 434
 Frank, H. S., 381, 382
 Frank, P., 418, 434
 Fraser, J. T., 832n.
 Freeman, I. M., 833n.
 Freudenthal, H., 128, 129n., 434
 Friedberg, R., 640, 641, 644
 Friedman, M., 843n.
 Fürth, R., 241, 242n., 257n., 434

 Gale, H. G., 355n., 438
 Gale, R., xxiii
 Galileo, 333, 406, 571
 Gal-Or, B., 831n.
 Gauss, J. K. F., 501, 502, 528, 781, 784
 Gerhardt, K., 153n., 434
 Gerlach, W., 399n., 434
 Geroch, R., 855n.
 Giannoni, C., 601, 602, 605, 606
 Ginzburg, V. L., 578n.
 Glymour, C., xxii, 556n., 754, 755, 764, 768, 772, 774, 782, 787, 788, 801, 807, 840, 842, 843, 853
 Gödel, K., 201, 202, 434, 824
 Goenner, H., 840
 Gold, T., 276, 278, 406, 407, 434
 Goldenburg, H. M., 578, 580, 629n.
 Goodhard, C. B., 57n., 434
 Goodier, J. N., 135n., 444
 Goodman, N., 190–191, 434, 480, 568
 Graves, J. C., 729n., 730, 747–750, 754, 756–764, 801, 855
 Griffiths, R. B., 664
 Grossman, N., 556n.
 Grünbaum, A., xvii, 94n., 215, 221n., 230n., 231n., 265–275, 278–280, 321–322, 354n., 368n., 369n., 397, 400n., 407n., 416n., 425n., 427n., 434–435, 436, 450, 567, 665n., 669, 678–679, 692–693, 701, 702, 707, 709, 726, 748, 752n., 754, 760, 762, 764, 767, 790, 797–799, 801f., 804, 829, 831, 832–833, 842n., 848–849; *Geometry and Chronometry in Philosophical Perspective*, xv, xvii, 450–452, 459, 486, 491, 495, 498, 510n., 534, 541, 542, 544, 545, 551, 552n., 556n., 567, 588, 597, 603–607, 620n., 621n., 683, 685n., 693n., 696n., 707, 731, 758, 762, 801, 805, 808, 840, 843, 845–846, 855; works on Zeno's paradoxes, 57n., 161, 168n., 179, 180, 208, 326n., 434, 435, 465, 552n., 567, 631n., 636n., 640, 645, 687, 707, 790, 797, 801, 804, 815–819, 822, 832, 849–851; works on falsifiability, 378n., 435, 456, 567, 588, 590, 608–609, 621n., 623n.
 Gulden, S., xv
 Gutting, G., 837

 Haber, F. C., 832n.
 Hadamard, J., 234, 332, 435
 Hall, D. B., 400n., 441
 Halmos, P. R., 166n., 435, 488n., 567
 Handler, P., 574n.
 Hanson, N. R., 292–299, 435
 Hardy, G. H., 172n., 435
 Hardy, L. H., 157n., 435
 Hartshorne, C., 166, 436
 Hasse, H., 160n., 436, 533, 567
 Havas, P., xv
 Hawking, S. W., 828n., 855n.
 Heath, T. L., 158n., 436
 Hempel, C. G., 290–311, 436, 449, 453n., 454–455, 506, 567, 718–722, 724–725, 727, 832, 837–838
 Hertz, H., 123, 436

- Hesse, M., 593, 598, 600, 601, 609, 610, 626
 Hessenberg, G., 772
 Heyting, A., 465, 567
 Hilbert, D., 150, 436
 Hill, E. L., xv, 265, 275
 Hille, E., 234, 436
 Hjelmlev, J., 176n, 436
 Hoagland, H., 57n, 436
 Hobson, E. W., 10n, 173n, 436, 469, 470, 567
 Hoffman, B., 735, 747, 751, 802
 Hölder, O., 127n, 331n, 334n, 436
 Holton, G., 342, 380, 436, 834, 839, 854
 Hood, P., 55n, 436
 Hooker, C. A., 841
 Hume, D., 372, 714
 Huntington, E. V., 162n, 189n, 215n, 436, 461, 468, 491, 512, 517, 564, 565, 567, 799, 802, 823n
 Hurewicz, W., 460, 484, 567
 Husserl, E., 152, 154
 Hutten, E. H., 105, 436
 Huyghens, C., 418
 Infeld, L., 714, 727
 Ingersoll, A. P., 629n
 Ives, H. E., 400, 710, 726
 Jaffe, B., 381, 436
 James, W., 208, 225n, 325, 436, 636n
 James & James Mathematics Dictionary, 463, 468n, 469, 470, 473, 533, 567, 656, 657, 665
 Jammer, M., 119n, 418, 422, 436, 833
 Jamis, A., xv, xxii, 225n, 450n, 536, 537, 561, 567, 629n, 655, 664, 669, 671, 707
 Jaynes, E. T., 646, 647, 664
 Jeffery, G. B., 702
 Jeffery, R. C., 624n
 Jeffreys, H., 73n, 436
 John, F., 233n, 437
 Jordan, P., 400, 423, 437
 Juhos, B., 572n
 Kaempffer, F. A., 421n, 437
 Kant, I., 153, 154, 179, 200, 201, 244, 268, 330, 332, 382, 550, 835
 Kaufmann, W., 373
 Kennard, E. H., 711, 712, 726
 Kennedy, R. E., 388, 399, 437, 721
 Kenyon, D. H., 834n
 Klein, F., 20n, 39, 100n, 121, 150, 151, 331n, 437
 Klein, O., 421n, 437
 Kline, M., xxii, 75, 768n, 772, 773, 783, 784, 802
 Kreisler, M. N., 822n
 Krips, H., 831, 832
 Kronecker, L., 647
 Kronheimer, E. H., 828n
 Kuerti, G., 383n, 443
 Kundt, W., 738, 747, 751, 802, 842
 Kuper, C. G., 822n
 Lacey, H. M., 823, 841
 Lakatos, I., 618, 849
 Lalande, J. F. F. de, 576
 Landau, L., 262n, 437, 464, 472, 567
 Lande, A., 220n, 250, 251, 268n, 437
 Langford, C. H., 215n, 437
 Laplace, P. S. de, 575
 Larmor, J., 371, 403, 408, 409, 725
 Laudan, L., xxiii, 586, 587, 629n
 Laugwitz, D., 761, 802
 Lauritsen, I., 711, 712, 726
 Lavoisier, A. L., 596
 Lechallas, G., 179, 330, 331, 437
 Leclercq, R., 132n, 437
 Lee, H. D. P., 167, 437
 Leftert, C. B., 854n
 Lefschetz, S., 163n, 437
 Leibniz, G. W., 159, 166, 179, 181, 188, 192, 197, 206, 418, 437, 547, 568, 729, 755
 Leonard, H. S., 480
 Leone, S. C., 383n, 443
 Leverrier, U. J. J., 575, 577
 Levi-Civita, T., 772, 784
 Lewin, K., 179, 437
 Lewis, C. L., 190n, 215n
 Lewis, D. K., 826
 Lie, S., 39, 40, 47, 64
 Lifschitz, E., 262n, 437, 464, 472, 567
 Lipschitz, R., 784
 Lipschutz, S., 462, 468n, 475, 529, 568
 Lipson, S. G., 822n
 Lorentz, H., 371, 380, 387, 409, 437, 718, 725, 824, 825, 827

- Loschmidt, J., 240, 248, 437
 Ludwig, G., 249n., 252, 438
 Luneburg, R. K., 154, 157, 438
 Luria, S., 158n., 160, 438, 532, 568, 810, 811
 Lyons, H., 55n., 400n., 438

 Mach, E., 16, 17, 40, 372, 382, 418, 425, 438, 714, 835, 839, 840
 Margenau, H., 14n., 134, 250n., 253n., 438
 Maritain, J., 147, 151, 438
 Marzke, R. F., 454, 503, 504, 551, 568, 735, 787, 802
 Massey, G. J., xvii, xxii, 449, 450, 460, 461, 465, 466, 477, 482, 487, 490, 493, 495, 508, 510n., 516, 519, 524, 532, 534, 540, 550, 554, 565, 568, 589, 758, 802, 809, 815, 816, 817, 818, 819, 820, 822, 847
 Maxwell, G., xv, 77, 138
 Maxwell, J. C., 233n., 438
 Mayr, E., 302n., 438
 McCall, S., xxiii
 McCaskey, S. W., 383n., 443
 McPhee, J., 678, 679, 682, 683
 Medawar, P. B., 571n.
 Mehlberg, H., xv, 179, 188, 206n., 211, 215n., 219, 220n., 258, 261, 265, 271, 275, 384, 385, 438, 681, 682, 707, 831, 856
 Menger, K., 161n., 438, 484, 568
 Merleau-Ponty, J., 843n.
 Mermin, N. D., 854n.
 Michelson, A. A., 355n., 379, 408, 438, 710, 836
 Milham, W. L., 55n., 438
 Miller, D. C., 383, 710
 Miller, S. L., 574n.
 Milne, E. A., 22, 23, 54, 55, 262n., 400, 410, 417, 438, 454n., 557, 560
 Minkowski, H., 191, 192, 326, 328
 Misner, C. W., 424n., 439, 728, 787, 855n.
 Möbius, A. F., 331n., 439
 Møller, C., 79n., 374n., 393, 399n., 438, 537, 538, 543, 544, 568, 605n.
 Morgenbesser, S., xv
 Morley, E. W., 710, 836
 Morrison, P., 850, 851
 Mostepanenko, A. M., 833
 Mostepanenko, V. M., 833
 Müller, G. H., 832n.

 Murphy, G. M., 14n., 438

 Nagel, E., xv, 91, 96, 117–120, 125, 127n., 439, 588n.
 Newcomb, S., 576
 Newcomb, W. A., xxiii
 Newton, I. (philosophical theory only), 4, 9, 12, 23, 26, 29, 47, 400, 418–421, 439, 535, 548–549, 561, 563, 568
 Newton, R. G., 687
 Noonan, T. W., 578n.
 Northrop, F. S. C., 6–8, 57n., 107–108, 134, 345n., 439

 Oparin, A. I., 571, 573, 574n.
 Oppenheim, P., 290–311, 436

 Page, L., 16n., 40n., 411n., 416–417, 439
 Palter, R. M., 63, 401, 425n., 427, 439
 Pannekoek, A., 576n., 577n.
 Panofsky, W., 388n., 439
 Parry, J. V. L., 400, 434
 Pascal, B., 159
 Pasteur, L., 571–574
 Pauli, W., 395n.
 Pearson, K., 502n., 783n.
 Peirce, C. S., 166
 Penrose, O., 277, 439
 Penrose, R., 828n., 855n.
 Pérard, A., 132n., 439
 Percival, I. C., 277, 439
 Perrett, W., 702
 Pfeffer, W., 57, 439
 Phillips, M., 388n., 439
 Pimenov, R. I., 828n.
 Pirani, F. A. E., 419n., 734–735, 736–737, 740–749, 756, 764, 792, 801, 802, 822n., 828
 Pittendrigh, C. S., 58n., 439
 Poincaré, H., 18–28, 44–56, 64, 85–94, 115–133, 156, 181n., 191, 227, 232, 241, 333–334, 370–371, 377, 408–409, 439–440, 548–549, 678, 680, 806–807
 Polanyi, M., 378–386, 440, 834–835, 836, 837, 839
 Popper, K. R., 44, 109n., 260–261, 264–277, 291–311, 440, 572n., 799, 831, 848–849
 Porter, J., xxii, 794n.
 Price, D. J., 55n., 440

- Putnam, H., xv, 15n., 31, 41, 113, 440, 449, 533, 563, 568, 588n., 620n., 731, 762, 802
- Quine, W. V. O., 106, 115, 131, 149, 440, 455, 506, 591
- Quinn, P., xxiii, 450, 587, 608, 610, 617n, 619n., 627, 629, 852, 853
- Rand, G., 157n., 435
- Read, J., 331n., 440
- Reichenbach, H., xv, xvii, 15, 104, 127, 128, 131, 133, 134, 187, 188n., 205, 208, 215, 222n., 241n., 278, 279, 288, 315, 318, 319, 322, 325, 330, 355n., 385, 417, 440, 441, 557, 613, 618, 619n., 699, 708, 748, 756, 760, 764, 795, 820, 826, 827, *Axiomatik der relativistischen Raum-Zeit-Lehre*, 179, 186, 201n., 341, 344n., 356n., 367n., 399n., 440, 543, 544, 568, 673n., 675, 707; *The Direction of Time*, 181, 188, 190n., 193n., 207n., 215, 218, 229n., 231n., 244n., 247n., 250n., 252, 253n., 254, 255, 261, 263, 288n., 313, 319, 322, 440, 789, 790, 798, 800, 829, 830; *The Philosophy of Space and Time*, xv, 30, 42, 44, 82, 105, 127, 152, 154, 180, 188, 193n., 201n., 222n., 316, 317, 330, 341, 353, 415n., 441, 471, 474, 552, 568, 666, 671, 673, 677, 678, 683, 689, 690, 691, 702, 704, 708, 763, 802, 853; *The Rise of Scientific Philosophy*, 89, 98, 104, 127, 382, 441, 835; works on the foundations of quantum mechanics, 181, 187, 250n., 252, 253n., 288, 292n., 319, 321, 441
- Reidemaster, K., 153n.
- Rescher, N., xv, xxiii, 292, 296, 441, 832
- Ricci, C., 784
- Richter, C. P., 57n., 441
- Richtmeyer, F. K., 711, 712, 726
- Riemann, B., xvii, xxii, 8, 18, 26, 28, 41, 43, 45, 48, 50, 56, 61, 64, 66, 115, 128, 129n., 419, 420, 425, 427, 441, 449, 451, 455, 460, 461, 471, 473, 476, 495, 504, 522, 526, 528, 534, 545, 549, 553, 568, 690, 691, 696, 698, 701, 703, 705, 707, 751, 755, 757, 760, 764, 772, 783, 784, 786, 788, 804, 805, 843, 848
- Rittler, M. C., 157n., 435
- Ritz, W., 375, 398n
- Robbins, H., 164n., 167n
- Robertson, H. P., 88n., 119, 369n., 399, 441, 575n., 806
- Romer, O., 355n., 699
- Rosentfeld, L., xviii, 248, 249
- Rosinger, K. F., 189n., 436
- Rossi, B., 400n., 441
- Rothstein, J., 288n., 441
- Rouquier, I., 131, 442
- Ruderman, M. A., 687
- Rudner, R., 292n., 311n
- Rush, J. H., 386n., 442
- Russell, B., 11n., 159n., 174, 175, 180, 442, 566, 740, 846, on the foundations of geometry, 12, 16, 44, 48, 127, 128, 442, 451, 458, 480, 497, 521, 548, 550, 556, 568, 805
- Saccheri, G., 311n., 442
- Sachs, R. G., 245n., 442, 797, 802, 856
- Sagnac, G., 355n
- Salecker, H., 151n., 442
- Salmon, W. C., xxii, xxiii, 292n., 561, 567, 618, 624, 629n., 666, 669 (co-author), 677, 695, 815, 832, 849, 850, 851, 854n.
- Sandage, A., 119n., 442
- Schaffner, K., xxiii, 854n
- Scheffler, I., 295, 442
- Schiff, L. I., 578n., 579n., 580
- Schiller, M., 537, 566, 770, 772, 793, 801, 840n.
- Schild, A., 734, 735, 736, 737, 740, 749, 756, 764, 792, 801, 802, 828
- Schiller, F., 385
- Schlick, M., 153n., 222n., 234, 236, 442
- Scholz, H., 160n., 179n., 436, 442, 533, 567
- Schorsch, E., 805
- Schouten, J. A., 770, 779, 782, 802
- Schrödinger, E., 245, 254, 258, 442, 504, 568, 787, 790
- Schwarzschild, K., 119, 420, 840
- Sciama, D. W., 421n., 442
- Scribner, C., 344n., 442, 702, 708
- Scriven, M., 292, 295, 299, 311, 442
- Sellars, W., xv, xxiii, 223n., 323n., 443, 450n.
- Shankland, R. S., 383n., 443, 710, 726, 836, 837

- Shepperd, J. A. H., 189n., 443
 Shimony, A., xv, 196
 Silberstein, L., 369n., 443, 727n., 840n.
 Sklar, L., 616n., 619n., 843n.
 Skyrms, F. B., 286
 Smart, J. J. C., xviii., 217, 235n., 289, 322n., 326n., 329, 443, 823, 831
 Sokolnikoff, I. S., 135n., 443
 Sommerfeld, A., 247n., 369n.
 Sommerville, D. M. Y., 126, 443
 Spiegel, E. A., 629n.
 Spinoza, B., 264
 Stachel, J., xxii, 374, 728n., 744, 746, 747, 749, 802
 Stanyukovic, K. P., 262n., 443
 Stebbing, L. S., 230n., 443
 Stein, H., 760, 761, 762, 802, 824, 852, 853
 Steinman, G., 574n.
 Stille, U., 132n., 443
 Stilwell, R., 400, 710
 Streater, R. F., 855
 Struik, D. J., 13n., 102n., 103n., 120n., 443
 Strükelberg, E., 207, 253n.
 Sudarshan, E. C. G., 822n., 827
 Suna, A., 146
 Suppes, P., xxii
 Swanson, J. W., 608, 616, 617n., 622–623n.
 Swinburne, R., 449, 550, 555, 559, 568
 Synge, J. L., 537, 541, 568, 664, 665, 731–735, 738, 740, 742, 745, 748, 749, 750, 802
 Tacquet, A., 158
 Tamm, I. E., 350n., 443
 Tannery, P., xviii., 167, 444
 Tarski, A., 47, 443, 463, 466, 565, 568
 Taub, A. H., 421, 443
 Taylor, G. I., 73n., 443
 Taylor, W. B., 182–185, 443
 Ter Haar, D., 241n., 444
 Terletsii, Y. P., 854n.
 Thomas, T. Y., 19n., 444
 Thomson, J. F., 631
 Thorndike, E. M., 388–399, 437, 721
 Timoshenko, S., 135n., 444
 Tolman, R. C., 208n., 240n., 241n., 247n., 256n., 376, 444, 577n., 681, 708, 727n., 830
 Torrance, C. C., 332n., 431
 Townes, C. H., 398n., 432
 Trocheris, M. G., 355n., 444
 Truesdell, C., 241n., 444
 Urey, H., 571, 573
 van Fraassen, B., xvii, xviii., xxiii, 450, 486, 491, 515–516, 518, 561, 567, 568, 669, 822–823, 828
 Veblen, O., 33n., 444, 462, 463n., 568
 Veronese, G., 173, 810, 813, 816, 818
 Vlastos, G., 632–633, 645
 Voigt, W., 408
 von Fritz, K., 176n.
 von Helmholtz, H., 128–129n., 152–156, 444
 von Laue, M., 369n., 374n., 400n., 409n., 444, 727n., 752, 802
 von Neumann, J., 250n., 252, 444, 662
 von Weizsäcker, C. F., 243–244, 288n., 444
 Waismann, F., 175n.
 Walker, A. G., 410n., 417
 Wallman, H., 460, 484, 567
 Ward, F. A. B., 55n.
 Watanabe, M. S., 234n., 248, 250n., 251–252, 263–264, 288n.
 Weatherburn, C. E., 464, 471–473, 540n., 568, 732n., 736, 784, 802
 Weber, H., 783, 802
 Weber, J., 369n.
 Weinberg, S., 840n., 855
 Weinstein, B., 132n.
 Weiss, P., 166, 208, 436
 Wertheimer, M., 710, 714, 716, 726, 838–839
 Weyl, H., xvii, 127n., 135, 176n., 181n., 326–329, 331n., 333, 336, 419n., 426, 495–497, 500–501, 503, 545, 551–552, 568, 630–631, 642, 644–645, 736–750, 751, 754, 756, 758, 759, 760, 761, 762, 764, 772, 802–803, 804, 842, 843, 845, 848
 Wheeler, J. A., 207, 253n., 423–424, 439, 450, 454, 467, 502–504, 547, 551, 568, 728–730, 735, 749–750, 753–755, 764–765, 767, 783, 786–788, 802, 803, 843
 Whitehead, A. N., 11n., 39n., 48–65, 208, 344–345, 425–428, 636n., 702

- Whitehead, J. H. C. , 33n , 444
 Whitrow, G. J. , 269n , 319, 321, 631, 636,
 637, 645, 833
 Whittaker, E. T. , 372, 376, 400, 409, 607n,
 727n
 Whyte, L. L. , 244, 245n , 414, 417
 Wiener, N. , 228, 230
 Wightman, A. S. , 855
 Wigner, E. P. , 151n , 442, 664n , 851
 Wilansky, A. , xv
 Willmore, T. G. , 735n , 803
 Winne, J. , xxii, 628, 629, 854
 Woodhouse, N. M. J. , 748, 746, 803
 Wright, J. P. , 202, 432, 824
 Yanase, M. M. , 280n
 Zahar, E. , 839
 Zeno of Elea, 158, 176, 180, 208, 532, 533,
 817, 818, 849, 850
 Zermelo, E. , 241
 Ziskel, E. , 186n , 269n

INDEX OF SUBJECTS

(Compiled by Theodore C. Falk)

- A. *See* Auxiliary assumptions collateral to a major hypothesis
- $\mathcal{A}$. *See* Affine structure
- Absoluteness:
 - of simultaneity, *see* Newtonian mechanics, indefinitely fast causal chains in; Quasi-Newtonian world
 - of space and time, *see* Metric, intrinsic; Relational theory of space-time
- Abstractionist account of theory-construction, 382, 835–836. *See also* Empiricism
- Accelerated frames in flat space-time. *See* Relativity theory, rotating disk in
- Acceleration:
 - dynamical, 68–75
 - secular (metrogenic), 70–73
- $\mathcal{A}$ -conventionality. *See* Conventionality, trivial semantic
- Action (human), 227, 232, 322, 825
- Additivity:
 - ϵ -additivity, 173–174, 810(def.) 813, 816–818
 - countable, 165, 171, 469(def.), 471, 476–478, 486–489, 531–535, 809–820, 846
 - finite, 165, 171, 469(def.), 471, 477–478, 487–489, 634, 810, 846
 - Massey (weakly) countable, 477(def.), 489, 813–817
- Ad hoc*, 710, 715–722, 724–725, 837–839
 - charged against Lorentz-Fitzgerald contraction hypothesis and Lorentz-Larmor time dilation hypothesis, 387–388, 392–394, 398n., 401, 709, 715–727, 838–839
 - charged against non-standard congruence definitions, 83, 89–91
 - charged against Zenonian 'addition' of zero lengths, 170–171
 - construed as the property of failing to entail significantly or interestingly different observational consequences, 719–721, 837–838
 - logical (systemic) vs. psychological construal, 387–388, 394, 718–721, 837–838
- Advance indicators, 283n., 286–287, 299, 309–311. *See also* Prediction
- Aether theory, 342, 380–382, 386–394, 398n., 400–409, 453, 603, 709, 715–717, 720–726, 824, 827, 836
- Affine structure ($\mathcal{A}$), affine coefficients, affine connection (Γ), 741–743, 765, 769–776, 779–780, 784, 787. *See also* Parallel transport of vectors
- Alternative metrization. *See* Metrizability, alternative
- Analytic-synthetic distinction, 108
- Angles:
 - measure of, in a Riemannian space, 19
 - of triangles, 102–103
 - of triangles, sums of, 109–110, 116, 124, 126, 807–808 (*see also* Stellar parallax)
 - See also* Congruence class for angles
- Anisotropy of time, xix, 180–191, 203, 206–208, 209–280, 314–317, 320, 347–349, 461, 564, 646, 746, 788–792, 796–800, 820–821, 828–832, 851
 - misleading potentialities of 'arrow' locution, 210, 222, 226, 769–797
 - nominal vs. physical specification of, 790, 792, 796–797, 798, 800
 - See also* Irreversibility
- Anthropomorphism, 728–730. *See also* Entropy, anthropomorphic vs. realist construal of
- Appearance (vs. reality), 617, 621, 622, 624
- Apriorism:
 - charged against conventionalism, 763
 - charged against philosophical appraisal of theories, 729, 754, 755
 - re* epistemology of scientific discovery, 379–383, 834, 836

- re* Euclideanism of physical space, 128, 130, 147, 151
- re* geometry of visual space, 152, 154
- re* inferences concerning the past, 244
- re* infinitude of space, 200, 201
- re* primacy of an intrinsic metric, 550, 551
- re* tri-dimensionality of space, 330
- 'Arrow' of time. *See* Anisotropy of time
- Assertibility of an explanandum, 293, 307
- Astronomy. *See* Clock, atomic vs astronomical; Clock, diurnal vs ephemeris; Precession of Mercury's perihelion; Solar oblateness, Stellar parallax
- Atomic:
 - clock, *see* Clock, atomic
 - constitution of matter, 237f., 644, 731, 850
 - space, *see* Space, discrete (granular)
- Auxiliary assumptions:
 - collateral to a major hypothesis, 107, 112, 114, 115, 136, 143, 383, 573, 575, 577, 581, 584, 586, 590, 602, 607, 610, 613, 629 (*see also* Duhemian thesis)
 - modifying a theory, *see* Lorentz-Fitzgerald hypothesis; Lorentz-Larmor hypothesis; *Ad hoc*
- Axiomatization. *See* Geometry, axiomatization of; Rational and empirical factors in axiomatization of physical theory; Relativity theory, axiomatization of
- Axioms:
 - for betweenness, 189n., 461, 463n.
 - for distance function, 163, 468
 - for measure function, 469
 - for spatial congruence, 11, 27, 29, 35, 36, 47, 52
- 'Backward' travel in time, 207
- Barometers, as advance indicators, 283n., 284, 286, 299, 307, 309, 311
- Bayes', theorem, 618
- B*-conventionality. *See* Conventionality, trivial semantic vs. nontrivial
- Becoming (the 'flow' of time), 5n., 56, 57n., 181, 208, 210, 217, 218, 224, 227, 251, 314, 329, 797, 832
- Betweenness, 96, 461, 465, 475, 477, 479, 481, 491, 493, 505, 522, 524, 564, 565
 - cyclic, 195, 196, 203, 799, 824
 - open (*o*-betweenness), 180(def.), 189, 193, 195, 196, 204, 214, 216, 279, 280, 348, 350, 823
 - separation-closure, 180, 189(def.), 193, 195, 196, 202, 203, 465, 823
 - temporal, 180, 181, 185, 186n., 188, 191, 193, 194(def.), 197, 206, 208, 209, 223, 230, 231, 685, 689, 700, 706, 791, 820, 822
- Bituration of nature, Whitehead's denial of, 11n., 48, 49, 57n., 427
- Binocular vision, 154, 157
- 'Block' universe, 327
- Bose-Einstein statistics, 203, 207, 268
- Boundary conditions, 180, 189, 202, 204, 205, 210, 217, 223, 233, 254, 258, 259, 263, 264, 272, 275, 278, 420, 422, 501, 545, 640, 645n., 831. *See also* Reversibility, de facto, Irreversibility, de facto, Contingent facts
- Bounded manifold. *See* Space, bounded
- Boyle's Law, 34, 38, 41
- Branch systems (temporarily quasi-closed systems), 181, 221, 233, 235, 236, 243, 254, 264, 277, 280, 281, 289, 320, 646, 658, 659, 665n., 789, 798, 800, 821, 829, 831
- Brownian motion, 261
- $\%.$ *See* Conformal structure
- Calculation, of an infinite sequence of digits, 631, 644, 645
- Cantor discontinuum. *See* Ternary set
- Cardinality, 160, 164, 174, 462, 529, 816, 817, 830
 - as basis for metric, 8, 10, 45, 476, 479, 485, 488, 490, 491, 495, 496, 500, 508, 509, 517, 526, 545, 550, 554, 559, 690, 843, 848
 - as an intrinsic property *vis-à-vis* equicardinality, 509, 529, 844, 848
 - metrics sensitive to, 484, 494, 555, 557
- Cauchy sum, 817, 818
- Causal:
 - chain, 28, 29, 182, 187, 189, 192, 203, 204, 206, 334, 347, 350, 369, 376, 670

- chain, genidentical, 182, 189(def.), 188–190, 193–196, 203, 204–207, 347–348, 350, 411, 537, 686–688
- chains faster than light, *see* Tachyons
- connectedness (' k -connectedness'), 187–188, 193–197, 207, 347–348, 822, 825
- connectibility, 28, 187–188, 194n., 203, 204, 315n., 319, 347–354, 366, 367, 370, 670, 686–687, 689, 692, 706, 713, 734, 827 (*see also* Light rays, as fastest causal chain; Simultaneity, topological)
- continuity, 180, 193, 194, 206, 785
- relation, asymmetric, 180, 181, 186, 188, 189–190, 218, 347–348, 745–746, 825–827
- relation, symmetric, 184n., 186, 188–191, 193, 205, 207–208, 213n., 216, 217, 348
- theory of time, 179–208, 223, 231, 345–346, 821–828
- Cause, 85, 303–313, 319, 321, 406–407
 - of accelerations, 68–77, 406
 - of Lorentz contractions, 404, 406, 408
- partial, 184n., 825
- ucceeding effect, *see* Retrocausality
- variational conception of being *the* cause, *see* Mark method
- Certainty, 301
- Ceteris paribus*, 191, 310, 311n.
- Change (vs. 'timelessness'), 327
- Chemical composition. *See* Deforming influences
- Christoffel symbols:
 - first kind, 768
 - second kind (metric connection), 769–772, 779, 792–793
- Chronometric method of metrizing space-time, 731–735, 740, 744–745, 748–751
- Chronometry, 12, 22–23, 50–60, 63, 66–80, 222n., 342–368, 666–708. *See also* Clock
- Clock, 52, 54–55, 193n., 400, 411–417, 453–455
 - atomic, 731–733, 735, 740, 749–751, 753
 - atomic (t) vs. astronomical (τ), 22–23, 55, 400, 410, 417, 454–455, 557–558, 560
 - atomic vs. gravitational, 735
 - axiom, 395–398
 - behavior under transport, 334, 346–347, 350n., 352, 367, 369–370, 683–707 (*see also* Clock, isochronous; Clocks, synchrony of, by slow transport of)
 - biological, 57–59
 - diurnal (T) vs. ephemeris (t , Newtonian time), 66–77, 122, 452–455, 553, 805
 - isochronous, as congruence standard, 66–67, 454, 545, 690 (*see also* Rods, rigid, and isochronous clocks)
 - paradox of STR, 854
 - retardation of STR, 361–362, 398n., 399–400, 710 (*see also* Lorentz-Larmor time dilation)
 - synchrony of, *see* Synchronization of clocks; Simultaneity, metrical
 - Clocks, synchrony of, by slow transport of, 667–671, 678, 683–707, 840, 851–854
 - Closed:
 - system, 186, 191, 208, 219–223, 230, 234–236, 239–247, 256–263, 280, 282, 285–289, 312, 322 (*see also* Branch systems)
 - time, 197–207, 272, 461, 465, 794, 799, 823–824 (*see also* Betweenness, separation closure)
 - Cluster concept. *See* Multiple criteria concept
 - Coarse-graining (in statistical mechanics). *See* Partitioning of phase space
 - Coherent initial conditions, 266–268, 270–271, 273, 275
 - Coincidence behavior. *See* Congruence standard; Rods
 - Collateral theory. *See* Auxiliary assumptions collateral to a major hypothesis
 - Color properties, 34–41, 458
 - Concordance:
 - of rod-coincidences, path-independent, *see* Riemann's Concordance Assumption
 - of metrical methods, *see* Metrical methods for space, time, concordance of
 - of clocks, *see* Clock, atomic vs. astronomical

- See also* Consistency
- Conditionals, subjunctive or counterfactual, 190, 546, 826
- Confirmation *See* Verification
- Conformal:
 structure (%), 739, 744, 746, 748, 751, 761, 764, 792, 794
 transformation, 19, 103, 736, 738, 794
- Congruence, 3, 4, 7, 11, 12, 23, 28, 32, 39, 44, 47, 61, 65, 130, 451, 452, 468, 469(def.), 471, 507, 550, 603, 607
 'definition' of, 14, 15, 43, 45, 61, 64, 84, 90, 92, 105, 119, 123, 127, 130, 133, 134, 142, 145, 152, 222, 415, 428, 548, 731, 795, 806 (*see also* Conventionality of congruence)
 'definition', epistemic status of, 130, 561
 standard for, 3, 11, 14, 42, 43, 44, 49, 60, 61, 87, 89, 139, 140, 426, 452, 455, 499, 501, 504, 510n., 546, 547, 548, 549, 553, 558, 559, 606, 748, 753, 764, 785, 787 (*see also* Clocks, Rods, Geodesic method of metrizing space-time; Chronometric method for metrizing space-time; Self-congruence)
 transformations, 39, 469, 470
- Congruence class (partition)
 for angles, 19, 101, 103, 120, 132, 735, 736
 for space(-time) intervals, 11, 12, 15, 16, 18, 19, 22, 27, 39, 78, 82, 87n., 98, 99, 109, 110, 131, 134, 140, 143, 452, 499, 507, 509, 542, 545, 547, 549, 550, 555, 557, 560, 562, 563, 735, 743, 842, 844, 845
 for time intervals, 4, 12, 16, 22, 23, 27, 28, 63, 78, 101, 122, 453, 455
- Congruences, incompatible, yielding same geodesics or same metric geometry, 14, 39, 54, 86, 87n., 98, 105, 735, 736, 767, 773. *See also* Metrizability, alternative
- Conscious organisms, 192, 197, 202, 208, 218, 251, 252, 318, 321, 329, 413, 415. *See also* Psychological time; Perceptual threshold for time awareness
- Conservation laws, 73, 74, 593, 598, 601, 653. *See also* CPT conservation
- Consistency
 of Euclidean and non-Euclidean geometry, 121, 149, 151
 of intrinsic curvature with intrinsic metric amorphousness, 764, 788
 of Cantor's theory of the interval, 172, 174, 818
 of metrical methods. *see* Metrical methods for space-time, consistency of
 presupposition of global time-orientation, 795, 796
 of transported-clock simultaneity, 670, 696, 705
 of transported-rod congruences. *see* Riemann's Concordance Assumption
- Contemplation of $\aleph_1$ subintervals. *See* Perceptual threshold for time awareness
- Context of discovery vs. context of justification. *See* Discovery, context of
- Contingent with respect to laws of nature (nomologically contingent facts), 180, 189, 205, 210, 211, 255, 259, 264, 272, 373, 640, 820, 821. *See also* Boundary conditions, Empiricism; Reversibility, de facto, Irreversibility, de facto
- Continuum. *See* Space, continuous; Numbers, real, Order, dense continuous
- Convenience. *See* Simplicity of theory
- Conventionality (and conventionalism):
 asserted of topological properties of space, 333, 337
 of congruence, 9n., 11, 12, 15, 32, 44, 46, 48, 50, 52, 53, 64, 127, 130, 144, 405, 412, 416, 449, 451, 456, 483, 495, 498, 500, 511, 520, 535, 683, 748, 756, 764 (*see also* Congruence 'definition')
 of congruence, spatial, 18, 23, 60, 62, 81, 85, 87, 90, 93, 94, 139, 147, 334, 395n., 680, 691, 696, 705
 of congruence, temporal, 23, 40, 53, 60, 63, 66, 72, 77
 of congruence, logical relation to alternative metrizability, 514, 515, 517, 548, 549, 559, 560, 561, 563

- of congruence, logical relation to
extrinsicity of the metric, 542,
556–560
- of direction in an extrinsic serial order,
209, 215
- geochronometric (GC), 28, 31–41, 75,
334, 410, 561
- linguistic, 115–131, 143–144, 806–807
- of metric relations other than congru-
ence, 507–508, 513–515, 519–521
- of simultaneity, 28–32, 35, 40, 61, 343,
349–355, 358–359, 367–368, 370, 377,
396, 397, 410–416, 451, 454, 456, 561,
666–671, 677, 679, 682–684, 688–693,
696–707, 712–715, 834, 852 (*see also*
Synchrony, standard and non-
standard)
- trivialized by model theory (MTT),
25–26, 32, 35, 46–47
- trivial semantic (TSC, *A*-conventionality)
vs. nontrivial (*B*-conventionality),
27–38, 75, 113, 451, 550, 758
- trivial vs. general, 692
- Coordinates, 505, 506, 681
 - Cartesian, 91, 96, 98–100, 161, 680
 - intrinsic, 504, 755
 - polar, 91, 100, 680, 778
 - spherical, 777, 780
 - time-orthogonal, 538–544, 604
 - for space, 87n., 96, 99–100, 209, 508,
522–523, 537, 777–780
 - for time, 209, 213n., 259–260, 279, 318,
349, 350n., 522–523, 670, 671,
689–692, 696–697, 700, 703–707, 791,
796, 805
 - for space-time, 16, 345, 504, 777–781
- Coordinative definition, 181, 222, 223,
387–388, 687–688, 706, 717–718,
760–762
- statistically based, 278–280
- See also* Congruence 'definition';
Conventionality
- Correctional laws, 43, 123–124, 132–134,
136, 139–142, 145, 149, 543, 546, 606,
731, 807. *See also* Deforming influ-
ences
- Cosmically local systems, 157, 246, 254,
256, 259–263, 272–273, 275, 277, 280,
798
- Cosmogony, 192, 198, 232, 406–407
- Cosmology, 22–23, 211, 221, 261, 275,
276–278, 417, 465, 537
- steady-state, 276–278
- See also* Entropy of the universe; Space-
time, infinitude of
- Counterfactuals. *See* Conditionals,
subjunctive
- Countable additivity. *See* Additivity,
countable
- Counting, 8, 496, 636–637, 844. *See also*
Cardinality, as basis for metric
- CPT-conservation, 797, 855–856
- Creation of the universe. *See* Cosmogony
- Crucial experiments, 570–571, 583–587
- CRUM. *See* Curvature, thesis of require-
ment of a unique metric
- Curvature:
 - Gaussian, 14, 20, 33, 97, 100, 102, 103,
146–147, 150–151, 428, 501, 528,
539n., 752, 767, 772, 777, 781, 783
 - Riemannian, 147, 150, 752, 772,
776–777, 781–782
 - Riemannn(-Christoffel) tensor,
covariant, 766–768, 770–785, 806
 - Riemann(-Christoffel) tensor, mixed,
766–768, 770–780, 784, 806
 - of space, variable, 90, 94–95, 147,
426–428, 473, 501–504, 750, 764–788,
806, 822, 841, 843
 - thesis of Curvature Requirement of a
Unique Metric (CRUM), 765–768,
772, 776, 782
 - thesis of feasibility of Metrical Un-
supportedness of Curvature (MUC),
765–788
- Cybernetics, 832
- Cyclic:
 - betweenness. *see* Betweenness, cyclic
 - theory of history, 198
- De facto. *See* Contingent facts; Boundary
conditions; Reversibility, de facto;
Irreversibility, de facto
- Definite metric. *See* Metric, definite and
indefinite
- Definition, coordinative. *See*
Coordinative definition; Congruence
definition

- Deforming (perturbational, distorting) influences, 3, 4, 12, 21, 36, 43, 44, 54n., 78, 81, 82, 85, 86, 91, 94, 104n., 109, 123, 124, 131, 132, 142, 145, 146, 149, 151, 186n., 404, 406, 408, 416, 455, 551, 612, 616, 618, 626, 684, 696, 731, 733, 747, 795. *See also* Correctional laws
- Degenerate intervals. *See* Points, Elements
- Dense order. *See* Order, dense
- Denumerable dense order. *See* Order, dense discontinuous; Space, denumerable dense
- Depth perception, 154, 157
- Determinism and indeterminism, 201, 202, 203, 210, 234, 248, 249, 272, 284, 286, 289, 299, 300, 312, 319, 325
- Deus ex machina*, 269, 271, 733, 734, 735, 742, 750
- Dicke's postulate (DP). *See* Gravitational oblateness
- Dictionary. *See* Intertranslatability, Coordinative definition
- Differential forces. *See* Deforming influences
- Diffusion equation, 233, 234, 247, 248
- Dimensional and dimensionless quantities, 771, 777, 783, 819
- Dimensionality of a space, 161, 330, 334, 460, 461, 820, 833, 834
- Discontinuity of a function of time, 637, 639, 643, 644, 850
- Discovery:
 context of, vs. context of justification (logical vs. psychological grounds), 150, 205, 231, 232, 381, 383, 386, 388, 392, 393, 394, 718, 725 (*see also* Relativity theory, special, history of; Methodological justifiability)
 of metric relations, *see* Measurement, epistemic role of; Metric, intrinsic
- Discrete order. *See* Order, discrete; Space, discrete (granular)
- Disequilibrium, 235, 254, 263, 283, 288, 289, 798, 829, 830. *See also* Equilibrium
- Distance function (Distance-metric, *D*-metric), 11, 12, 14, 18, 19, 33, 87, 88, 96, 105, 163, 459, 468(def.), 475, 483, 493, 499, 500, 505, 507, 510, 512, 515, 518, 522, 529, 542, 544, 547, 548, 554, 556, 560, 561, 664
- Diurnal time (*T*). *See* Clock, diurnal vs. ephemeris
- D*-metric. *See* Distance function
- Doppler effect, 409, 710, 724
- D*-thesis. *See* Duhemian thesis
- Dualism, mind-body, 328. *See also* Mind-dependence of becoming
- Duhemian thesis (*D*-thesis), 108, 115, 123, 126, 131, 147, 149, 586, 629, 807, 808, 839, 848
 charge of triviality against, 111, 114, 588, 593, 598, 601, 607, 610
 subtheses D1 and D2 articulated, 587
- Dyadic relation. *See* Relational properties; Cardinality, *vis-à-vis* equicardinality
- Dynamics, 639, 640, 643. *See also* Cause, of accelerations
- e*. *See* Synchrony, standard and non-standard
- Ego. *See* Conscious organism
- Ehrenfest curve, 241, 242, 251
- Einstein's field equations. *See* Relativity theory, general
- Electrodynamics, XIX, 240, 250, 372, 373, 374, 379, 380, 620n., 728, 751, 787, 789, 842
- Electromagnetic causal chain. *See* Light rays
- Elements of a manifold, 9, 158, 176, 458, 464, 496, 808, 809, 817. *See also* Points; Instants; Events
- Embarrassing result (*E*-result), 387, 716, 718, 720. *See also* Evidence, disconfirmatory
- Embedding space. *See* Hyperspace
- Empiricism:
 crude, *see* Abstractionism
 re Duhemian thesis, 110, 114
 re geometry, 45, 127, 131, 135, 142, 152, 154
 re topology of space and time, 201, 330, 337
 See also Testability; Rationalism and empiricism in physical theory

- Empty curved space-time. *See* Geometrodynamics (GMD)
- Entailment, 506, 607
- Entropy, 181, 208, 216, 219, 230, 234, 247, 254, 264, 265, 289, 290, 646, 665, 754, 788, 791, 798, 800, 821, 828, 831, 851
 anthropomorphic vs. realist construal of, 646, 648, 659, 663, 664, 754, 829, 851
 law, *see* Thermodynamics, second law of
 as measure of degree of homogeneity or disorder, 237, 239, 650, 659, 663
 as measure of number of underlying microstates, 237, 239, 262, 647, 662
 as measure of probability of occurrence, 237, 239, 650, 659, 661, 664
 of the universe, 244, 261, 263, 267, 268
See also Equilibrium; Disequilibrium; Homogeneity; Partitioning of phase space, dependence of entropy on
- Ephemeris time (*t*). *See* Clock, diurnal vs. ephemeris
- Epistemic considerations. *See* Assertibility of an explanandum; Congruence 'definition', epistemic status of; Conscious organisms; Discovery, context of; Measurement, epistemic role of; Methodological justifiability; Metric, extrinsic, epistemic uncertainty regarding its extrinsicity; Microstates, ignorance of; Prediction, epistemic failure of in a probabilistic world; Retrodiction, epistemic asymmetry with prediction
- Equality relation. *See* Equivalence relation
- Equilibrium, 216, 223, 231, 239, 258, 262, 290, 660, 829, 830. *See also* Homogeneity
- Equivalence:
 of descriptions, 41, 68, 75, 89, 119, 417, 453n., 557, 558, 806, 807
 principle of, for inertial and gravitational mass, 91, 746, 748, 763, 764
 relation, 11, 23, 27, 28, 32, 39, 47, 130, 452, 468, 494, 506, 509, 513, 515, 519, 520, 524, 525, 526, 530, 539, 542, 545, 548 (*see also* Congruence)
 relation, intensional construal of, 506, 507, 509, 513, 514, 524
- Essence, 152
- Ether theory. *See* Aether theory
- Euclidean geometry. *See* Geometry, Euclidean
- Events:
 as fundamental constituents of the universe, *see* Relational theory of space-time
 individuation of, *see* Individuation
 punctal, 458, 536, 537
- Evidence, disconfirmatory, 106, 112, 114, 137, 142, 383. *See also* Falsifiability; Embarrassing result
- Evolution, biological, 192, 299, 303, 306, 628
- Expansion:
 of the universe (astronomical red shift), 276, 278, 539, 541
 universal nocturnal (philosophical *Gedankenexperiment*), 42, 43
- 'Experimenticism'. *See* Abstractionism
- Explanation, 290, 313, 587, 589, 608, 610, 832
- Explanatory parity, 66, 67, 68, 72
- Extension, spatio-temporal, 476, 477, 481, 483, 485, 495. *See also* Zeno's paradoxes, of metrical extension
- Extensionality (vs. intensionality), 451, 505, 507, 509
- Extensionless elements. *See* Elements; Events; Points; Instants; Zeno's paradoxes, of metrical extension
- Extreme elements. *See* Space, bounded
- Extrinsic feature of a manifold, 460, 462, 465, 506(def.), 507, 510n., 563, 566, 752, 844, 848. *See also* Metric, extrinsic
- Factlike. *See* Contingent facts
- Falsifiability of hypotheses, 106, 115, 131, 132, 135, 144, 389, 391, 393, 449, 450, 456, 569, 629, 848, 849. *See also* Testability; Verifiability; Empiricism; Crucial experiment
- Finsler space and Finslerian metric, 732, 733, 735, 742, 745
- Flat space-times. *See* Geometry, Euclidean; Space-time, Minkowskian
- 'Flow' of time. *See* Becoming

- Fluctuation phenomena. *See* Statistical mechanics
- Footprints, as retrodictive indicators, 234, 236, 281, 282, 285, 829
- Framework principle, 335, 336
- Freely falling geodesic test particles
 - massive (as time-like geodesics), 546, 734, 737, 739, 744, 746, 751, 753, 756, 764, 841, 842
 - massless, *see* Light rays, as null geodesics
- Free mobility, axiom of, 95, 129
- Free will, 321, 322, 833
- Future. *See* 'Now', Prediction, Advance indicators, Becoming
- Galilean:
 - frames, *see* Inertial frames
 - relativity, *see* Newtonian mechanics
- Gas, (statistical) mechanics of. *See* Maxwell-Boltzmann gas statistics
- Gaussian curvature. *See* Curvature, Gaussian
- GC. *See* Conventionality, geochronometric
- General property. *See* Property, general
- General theory of relativity (GTR). *See* Relativity theory, general
- Genidentity. *See* Causal chain, genidentical
- Geo-chronometric conventionalism (GC). *See* Conventionality, geochronometric
- Geodesic, 13, 14, 19, 21, 23, 97, 101, 103, 109, 110, 116, 119, 121, 124, 201, 546, 807, 808, 824
 - mapping, 736 (*see also* Projective geometry)
 - method of metrizing space-time, 546, 734, 751, 756, 764, 842
 - null, *see* Light rays as null geodesics
 - time-like, *see* Freely falling particles, massive
- See also* Straight line
- Geometrodynamics (GMD), 423, 450, 467, 500, 504, 536, 547, 728, 730, 734, 749, 765, 782, 783, 788, 843
- Geometry:
 - absolute, 138, 157n.
 - analytic, 150, 159, 160
 - axiomatization of, 130, 137, 138, 150, 156, 308
 - differential, 121, 139, 159, 160 (*see also* Metric tensor)
 - Euclidean, 18, 21, 24, 26, 32, 33, 39, 48, 54n., 63, 83, 86, 91, 92, 97, 98, 104, 109, 115, 130, 133, 151, 154, 157, 163, 201, 331, 470, 475, 802, 535, 539n., 544, 546, 581, 608, 611, 615, 623, 627, 767, 785, 806, 808
 - hyperbolic, 20, 21, 23, 85, 97, 109, 115, 116, 121, 138, 150, 154, 157, 544, 605, 808
 - metric, *see* Metric geometry
 - non-Euclidean, 24, 26, 32, 33, 39, 48, 63, 92, 100, 109, 116, 126, 129n., 139, 141, 145, 148, 150, 154, 156, 201, 806, 807 (*see also* Curvature of space)
 - spherical, 86n., 97, 103, 109, 115, 138
 - synthetic, 33, 119, 121, 137, 806, 807
- GMD. *See* Geometrodynamics
- Granular space. *See* Space, discrete (granular)
- Gravitational
 - collapse, 855
 - field, 79, 80, 90, 94, 121, 123, 421n., 546, 620n., 670, 684, 686, 689, 728, 731, 733, 840, 842
 - mass, *see* Equivalence principle for gravitational and inertial mass
 - multipolarity, 746, 748, 757
 - oblateness, Dicke's postulate (DP) of, 581, 584, 586
- GRCS. *See* Synchronism, generalized rule of committed
- Group theory, 39, 47, 64, 65, 234n.
- GTR. *See* Relativity theory, general
- H*-explanation and *H*-prediction, 291, 300, 303
- Holism. *See* Duhemian thesis; Theoretical system
- Homeomorphisms, 529, 530, 820, 821
- Homocentricity. *See* Operationism
- Homogeneity:
 - macroscopic, 237, 650, 659, 661, 662, 663 (*see also* Entropy; Equilibrium)
 - with respect to metrical extension of singletons, 458, 459, 462, 463, 466, 467, 479, 485, 486, 499, 501, 515, 535, 536, 547, 555, 755

- H*-theorem of Boltzmann, 239–247
 Huyghens's principle, 332–333
 Hyperbolic geometry. *See* Geometry, hyperbolic
 Hyperspace, embedding, 149–151, 202, 326n., 329, 330–332, 501–502, 537
 Hypotheses. *See* Falsifiability; Testability; Verification; Evidence; Auxiliary assumptions; *Ad hoc*
 Idealism, 251–252
 Identity:
 crypto-, 621–623
 metrics sensitive to, 487, 494–495
 of indiscernibles, 197–206, 547, 755
I-extrinsic(-intrinsic) metric. *See* Metric, extrinsic (intrinsic) to the intervals of a manifold
 IMA. *See* Metric amorphousness, intrinsic
I-manifold. *See* Inhomogeneous manifold
 Implosion, as temporal inverse of explosion, 266–271
 Indefinite metric. *See* Metric, definite and indefinite
 Indeterminism. *See* Determinism and indeterminism
 Indicator:
 of past, *see* Traces
 of future, *see* Advance indicators
 Individuals, 480, 504, 505. *See also* Elements of a manifold
 Individuation:
 of events, 204, 205
 of light rays, 745
 of molecules, 238
 of space-time points, 6, 345–346, 450, 458, 466–467, 493, 503–504, 535–536, 547, 755, 843 (*see also* Relational theory of space-time)
 See also Identity of indiscernibles
 Induction, 137, 617, 621–626, 646, 711, 848–849. *See also* Simplicity, inductive; Duhemian thesis
 Inertia, 101, 418–419, 422. *See also* Mass
 Inertial frames, 414–415, 420, 452–453, 537, 544, 604–607, 664, 678–683, 720, 805–806, 854. *See also* Freely falling particles
 Inferability, 293–307
 Infinite divisibility. *See* Order, dense; Zeno's paradoxes
 Infinite series, 168, 173, 630–644, 815–817
 'last' term of, 166–168, 633, 634–636, 637, 639, 640, 643, 815
 Infinitesimals, 166–167
 Infinitude of space. *See* Space(-time), infinitude of
 Infinity machines, 630–645, 849–850
 Influence chains. *See* Causal chains
 Information. *See* Trace
 Inhomogeneous manifold (*I*-manifold), 459–460(def.), 462, 466, 467, 479, 498
 Initial conditions. *See* Boundary conditions; Coherent initial conditions
 Instants, 458, 479–483, 536, 690. *See also* Elements of a manifold
 Insufficient reason (Laplacian equi-ignorance or indifference) 486, 662
 Intensionality. *See* Relations, intensional construal of; Equivalence relation, intensional construal of
 Intentions, as advance indicators, 286–287, 322
 Interaction, 243, 251n., 255–257, 263–264, 282–289, 296, 298, 300, 302, 312, 829
 See also Trace
 Interpretations of formal calculi, 47, 117–118, 127, 130, 132, 134, 137, 143, 174, 387, 512, 717. *See also* Coordinative definition
 Intertranslatability, 121, 131
 Intervals, 163–173, 463(def.), 460–558
 passim, 611, 808–818, 844–848
 quasi-disjoint, 528, 847
 See also Congruence class for intervals; Metric, intrinsic to intervals
 Intrinsic feature of a manifold, 460–462, 465, 467–468, 483, 501–502, 504, 506(def.)–509, 510n., 513–515, 519–520, 524, 539, 557, 562–563, 564, 690, 752, 760. *See also* Metric, intrinsic
 Intrinsic metric amorphousness (IMA). *See* Metric amorphousness, intrinsic
 Intuition, 149–150, 153n., 154, 156–157, 198, 202, 384, 660–661. *See also* Perceptualism; Apriorism

- Intuitionistic mathematics, 337
 Irreversibility and unreversedness,
 183, 190, 201n, 206, 209, 220, 234,
 236, 314, 461, 564, 789, 791, 832,
 strong (nomological), 210, 211, 216,
 272, 273, 791, 796, 797, 800
 weak (de facto, nomologically contin-
 gent), 206, 210, 211, 216, 219,
 254, 278, 791, 796, 800
 See also Anisotropy of time
 Isochronism. *See* Clock, isochronous.
 Rod, rigid, and isochronous clock
 Isotropy:
 of space, 355
 of time, 209, 212, 214, 216, 217, 232,
 272 (*see also* Anisotropy of time)
I-triplets, 477, 478, 528

 Jacobian, 331, 332
 Justification. *See* Discovery context of, vs.
 context of justification, Methodologi-
 cal justifiability

 Kelvin's thermodynamic scale of tempera-
 ture. *See* Thermometrics
k-connectedness. *See* Causal connected-
 ness
 Kennedy-Thorndike experiment, 388, 393,
 395, 398, 399, 721, 724
 Kinematics, 70, 73, 630, 645. *See also*
 Light rays as means of signal syn-
 chrony
K-interval, 525(def.)

 Laplacian demon, 146, 556
 Law of nature, 210, 211, 272, 275, 278,
 820. *See also* Irreversibility, strong;
 Reversibility, strong
 Law-like sentence, 506
 Learning, 155, 156
 Lebesgue integration, 166, 819
 Legato run. *See* Z-run, legato
 Length function (*L*-metric), 160, 166,
 171, 173, 470, 478, 483, 488, 493, 499,
 505, 507, 515, 538, 542, 558, 611, 614,
 679, 680, 681, 721, 732, 741, 742, 744,
 771, 777, 782, 808, 820, 844
 Light rays:
 as fastest causal chains *in vacuo*, 191,
 204, 334, 342, 348, 380, 386, 366, 367,
 369, 377, 399, 410, 640, 667, 670, 684,
 686, 689, 692, 699, 700, 713, 715, 737,
 834. *See also* Tachyons, Newtonian
 mechanics, indefinitely fast causal
 chains-in)
 as gemidentical objects, 182, 203
 as means for signal synchrony of clocks,
 410, 417, 482, 483, 602, 607, 666, 683,
 696, 707, 881, 882
 as (null) geodesics (generators of null
 cones), 109, 110, 115, 125, 135,
 472, 473, 492, 846, 733, 749, 751, 753,
 756, 761, 764, 789, 792, 795, 799,
 806, 808, 841
 'topological' condition for. *See* 'Topo-
 logical' condition for light
 Light, theory of:
 corpuscular, 274, 275, 585
 emission, 386n, 395n
 undulatory, 274, 275
 Light velocity
 causal chains faster than, *see* Tachyons
 constancy of (the '[two-way] light
 principle'), 79, 143, 342, 343, 370,
 375, 378, 386, 387, 394, 399, 401, 417,
 453, 602, 671, 676, 680, 698, 703,
 711, 715, 726, 854
 limiting character of, *see* Light rays as
 fastest causal chains
 one way, 354, 367, 370, 377, 386n., 397,
 672, 673, 698, 704, 712
 Light waves, 'stationary' (Einstein's
 Gedankenexperiment), 372, 374
 Lightning bolt experiment of Einstein,
 357, 359, 367
 Limits, theory of, 165, 172, 173, 635, 636,
 816, 817
L-metric. *See* Length function
 Lobachevskian geometry. *See* Geometry,
 hyperbolic
 Local. *See* cosmologically local
 Logic:
 non-standard, 112, 114
 of scientific discovery, *see* Methodo-
 logical justifiability
 Lorentz-Fitzgerald contraction hypothesis,
 342, 387, 394, 398, 401, 406, 715, 718,
 720, 727, 838, 839. *See also* *Ad hoc*

- Lorentz-Larmor time dilation hypothesis, 391–393, 398n., 403–404, 667, 723–726. *See also Ad hoc*
- Lorentz transformations:
 interpreted aether-theoretically, 401–404, 408–409, 723, 724, 824
 interpreted relativistically, 361–366, 375, 376, 399–404, 408–409, 668, 671, 678, 680, 682, 687, 693, 715, 724, 806
- Mach's Principle, 9, 418–424, 425, 833, 840
- Macro-state (in statistical mechanics), 205, 217, 224n., 237–240, 244n., 246, 248–249, 256, 646–664, 830
- Manifolds, 8–9, 153n., 458–461, 497, 509
 multiply extended, 460–461, 463–464, 475, 489, 490–491
 singly extended, 460–468, 469, 475, 491, 518
 See also Space; Elements of a manifold; Inhomogeneous manifold; Numbers
- Manna. *See Deus ex machina*
- Mark method of Reichenbach, 180–189, 208, 826–827
- Mass, 16–18, 40–41, 91, 418–424, 593, 599–601, 691–692, 746. *See also*
 Equivalence principle, for gravitational and inertial mass
- Massive test particles. *See* Freely falling particles, massive
- Massless test particles. *See* Light rays as null geodesics
- Maxwell-Boltzmann gas statistics, 205, 206, 238, 262, 267
- Maxwell's:
 demon, 288
 equations, *see* Electrodynamics
- Meaning (of words), 24–38, 47, 111, 113, 140, 550, 590–607
- Measure:
 function (Measure-metric, *M*-metric), 468, 469(def.)–471, 474–495, 499, 507–510, 512–517, 519–522, 524–527, 531–535, 540, 547–548, 554–556, 809, 810, 819, 845–846
 theory, 10, 160, 166, 172, 810, 851
- Measurement, epistemic role of, 6, 45–46, 352, 368, 401, 548, 561, 689, 748, 753, 785, 787. *See also* Deforming influences; Metric; Congruence; Metrogenic
- Mechanism vs. teleology, 281, 311–313
- Memory, 199, 212, 225–226, 263, 282–283, 289–290, 325. *See also* Trace; Conscious organisms
- Mercury. *See* Precession of Mercury's perihelion
- Metaphorical locutions. *See* Anisotropy of time, misleading potentialities of 'arrow' locution; 'Universal forces'
- Metarelativity. *See* Tachyons
- Meter stick. *See* Rigid rod; Standard meter
- Methodological justifiability of scientific hypotheses, 377–385, 626, 834–837, 848–849. *See also Ad hoc*
- Metric:
 alternative, *see* Metrizable, alternative
 amorphousness, intrinsic (IMA), 6, 10, 11, 16, 27, 119, 153, 334, 455, 495, 498–499, 514, 526–528, 535, 547–550, 553–555, 559, 563, 774, 786–788 (*see also* Riemann's Metrical Hypothesis [RMH]); Conventionality of congruence
 cardinality-based, *see* Cardinality as basis for metric
 coefficients, 19–20, 86, 87n., 97, 99–100, 420, 471, 472, 501, 538, 542, 545, 738, 777–781 (*see also* Metric tensor)
 components of (equality, ordinal, ratio, unit), 507–508, 513, 519–521, 526
 connection, *see* Christoffel symbol, second kind
 definite and indefinite, 472–474, 476, 489, 492, 527, 559, 732, 736, 757
 distance (*D*-), *see* Distance function
 extrinsic, xvii, 8–9, 11, 45, 48, 81, 129n., 426, 449–451, 455, 465, 495–563, 751–755, 758, 804
 extrinsic, epistemic uncertainty regarding its extrinsicity, 509–510, 512, 523, 530–531, 565
 extrinsic to the intervals of a manifold (*I*-extrinsic), 507(def.), 510
 extrinsic to a manifold (*S*-extrinsic), 510(def.), 527–531, 540–542, 545, 556, 562

- geometry, 10, 12, 14, 26, 63, 64, 81, 105, 335, 336, 419, 424, 426, 428, 498, 528, 549, 610, 615, 624, 627, 663, 730, 788
 homogeneity of singletons, *see* Homogeneity with respect to metrical extension
 intrinsic, xvii, 4, 12, 16, 17, 23, 26, 28, 33, 38, 39, 42, 43, 44, 47, 48, 56, 60, 63, 84, 129n., 345, 401, 404, 409, 418, 422, 449, 451, 461, 465, 495, 565, 751, 760, 785, 788, 804, 843, 848
 intrinsic to the intervals of a manifold (*I*-intrinsic), 507(def.) 510, 512, 522, 524, 527, 531, 532, 535, 539, 541, 545, 547, 548, 554, 556
 intrinsic, lack of, *see* Metric amorphousness
 intrinsic to a manifold (*S*-intrinsic), 509(def.) 510, 511, 512, 513, 517, 518, 520, 521, 522, 526, 527, 530, 539, 541, 545, 547, 554, 555, 560
 intrinsic to the ordered pairs of a manifold (*Op*-intrinsic), 508, 511, 513, 515, 518, 521, 522, 523, 525, 526, 527, 547
 length (*L*-), *see* Length function
 measure (*M*-), *see* Measure function
 nontriviality of, 451, 455, 459, 474, 495, 498, 501, 512, 515, 535, 548, 550, 553, 559, 844, 846, 848
 nontriviality of, criteria for, 474, 495
 'pathological', 458, 487, 488, 490, 493, 494, 495, 508, 519, 520, 524, 532, 554
 Riemann (*R*-), *see* Riemann-metric spaces, 162, 163, 468, 473, 750, 788, 789, 790, 792, 796, 799, 800
 tensor q_{ik} , 11, 14, 33, 82, 86, 89, 97, 105, 131, 139, 151, 464, 535, 538, 542, 620n., 736, 737, 740, 744, 745, 751, 757, 762, 765, 781, 784, 789, 792, 794 (*see also* Metric coefficients)
 trivial, *see* Metric, nontriviality of
 Unsupportedness of intrinsic Curvature (MUC), *see* Curvature, metrically unsupported
See also Congruence; Geometry; Space-time
 Metrical field, dynamical conception of, 56, 419, 423, 501
 connection with Riemann's Metrical Hypothesis (RMH), 9, 419, 420, 425, 427, 500, 501, 545, 546
See also Mach's principle; Relativity theory, general; dynamically-based metric of
 Metrical methods for space-time concordance of, 750, 751
 epistemic circularity of, 733, 745, 749
See also Chronometric method; Geodesic method; Rods, rigid; Clock, isochronous
 Metrizable, alternative, 12n., 23, 27, 32, 36, 41, 46, 53, 63, 64, 495, 514, 517, 522, 523, 547, 549, 559, 563
 for space-time intervals, 14, 15, 18, 23, 84, 96, 105, 115, 126, 129n., 142, 147, 428, 544, 777, 805
 for simultaneity, *see* Synchrony; standard and non-standard
 for time intervals, 22, 23, 66, 80, 118, 805
 nontrivial construal of, 514, 547
 normative construal of, 548, 549, 560, 561, 563, 842
See also Metric amorphousness
 Metrogenic
 acceleration, *see* Acceleration, secular
 contraction, 402, 840
 irreversibility in quantum mechanics, 249, 253
See also Measurement
 Michelson-Morley experiment:
 historical role of, 379, 383, 710, 835, 837
 physical significance of, 386, 390, 393, 395, 397, 399, 401, 406, 603, 604, 607, 710, 712, 716, 718, 720, 724, 837, 839
 Micro-state (in statistical mechanics), 205, 206, 217, 224n., 237, 241, 244n., 248, 256, 257, 646, 649, 664, 830, 831
 epistemic role of ignorance of, 238n., 245, 646, 648, 662, 664
 Mind-dependence of becoming, 323, 329, 832. *See also* Becoming; Conscious organisms
Minima perceptibilia. *See* Perceptual threshold
 Minkowski space-time. *See* Space-time, Minkowskian

- Miracles, 406–407
- Mixing processes, 211, 225, 227–228, 240, 254, 267
- M*-metric. *See* Measure function
- Möbius strip, 331
- Model:
 theory. *see* Interpretations of formal calculi; Conventionality, trivialized by model theory
 universes, 197–206
- MUC. *See* Curvature, thesis of metric unsupportedness
- Multiple criteria (cluster) concept, 14–15, 62, 134, 278, 387–388, 451–453, 455
- Multipolarity. *See* Gravitational multipolarity
- Naming, in the geometry of space-time.
 See Conventionality, linguistic; Conventionality, trivialized by semantics; Anisotropy of time, nominal vs. physical specification of
- Natural numbers. *See* Numbers, natural
- Naturalism, 192
- N*-chain, 194(def.)–196
- Necessity of an event for the *k*-connectedness of two other events, 194, 822
- Neptune, discovery of, 576
- Newtonian mechanics:
 continuous space of, 501
 deterministic character of, 299
 gravitation in, 94
 indefinitely fast causal chains in, 343, 346, 349, 351, 367–368, 370, 456, 670, 684–689, 691–692, 702, 712, 791, 834
 relation to relativity theory, 123, 343, 370, 408, 575–579, 598–601, 834
 time metric of, 49–56, 63, 66–77, 100–101, 118, 204, 453–455 (*see also* Clock, diurnal vs. ephemeris)
 time-symmetry of, 213–214, 216, 239–240, 312, 790–791
 See also Cause, of accelerations
- Nominal specification, 744. *See also* Anisotropy of time, nominal vs physical specification of; Conventionality, linguistic
- Nomic regularity. *See* Law
- Nomologically contingent facts. *See* Contingent facts
- Non-Euclidean geometry. *See* Geometry, non-Euclidean
- Non-standard:
 congruence. *see* Metrizable, alternative
 logic. *see* Logic, non-standard
 synchrony. *see* Synchrony, non-standard
- Nontriviality. *See* Conventionality, trivial; Duhemian thesis, charge of triviality against; Metric, nontrivial; Metrizable, alternative, nontrivial construal of
- Normative alternative metrizable. *See* Metrizable, alternative, normative
- Norms, epistemological, in the history of science, 835, 839
- 'Now', 181, 210, 217–218, 223, 226, 291, 314–326, 329. *See also* Becoming
- N*-quadruplet, 194(def.)–196
- Null cone. *See* Light rays as null geodesics
- Numbers:
 imaginary, 471, 472
 natural, 458, 460, 462, 467, 478–480, 517–518, 566, 630
 pure. *see* Dimensional and dimensionless quantities
 rational, 162, 175, 458, 460, 462, 467, 516–517
 real, 16–17, 150, 162, 175, 215, 219, 223, 458–460, 462, 467–468, 498–499, 505, 512–516, 517, 518, 556, 810
 real, as space-time coordinates, 16, 202, 209, 216, 259–260, 279
 See also Order, types of; Cardinality
- o*-betweenness. *See* Betweenness, open
- Observation, 136–137, 184–185, 193n., 197, 199, 251–253, 321, 546, 719.
 See also Falsifiability; Testability; Evidence; Measurement
- Occam's razor. *See* Simplicity of theory
- One-to-one correspondence, 844–848
- Ontology, xvii, 506, 527, 547, 557, 561, 728–788, 792, 800, 824–827, 829, 841, 843, 851

- Open:
 betweenness, *see* Betweenness, open
 cluster concept, *see* Multiple criteria concept
 system, *see* Interaction, Space-time
 infinitude of
 texture, *see* Multiple criteria concept
 Theoretical terms
 time, 197–208, 217, 218, 223, 272
 Operational definition. *See* Coordinative definition
 Operationism, 15, 41–44, 61–62, 119, 134, 174, 175–176, 230–232, 368, 412
 Op-intrinsic metric. *See* Metric, intrinsic to the ordered pairs of a manifold
 Optics, geometrical, 109–110, 115, 124–126, 274, 585–586, 597, 603–606, 806–808
 Order, types of
 dense, 11, 160, 167–170, 174–176, 180–208, 463, 636n, 816
 dense continuous, 8–11, 12n, 16, 158–176, 180, 194, 334–337, 820–850
 dense discontinuous, 8n, 9–10, 162, 164, 165–166, 171, 172, 175, 267, 462–463, 809
 discrete, 9, 10
 See also Number, Betweenness, Space times
 Ordinary language, 185, 198, 307–308, 356, 637n, 825
 philosophy, 35n, 223, 294, 351–352, 550
 $\mathcal{O}$. *See* Projective geometry
 Pair-production and pair-annihilation, 207, 253n
 Parallel:
 postulate, 138
 transport (displacement) of vectors, 743, 789, 793–796
 Paresis, as test case for symmetry of explanation and prediction, 299, 303–306
 Particle, 207, 728
 physics, 797–798, 855–856
 Partitioning:
 into equivalence classes, *see* Equivalence class; Congruence class
 of phase space (in statistical mechanics), 208–237, 238
 of phase space, dependence of entropy on, 646–664
 Passage of time. *See* Becoming
 Past. *See* Now–Retrodiction, Traces, Anisotropy of time
 Path independence and path-dependence:
 of congruence. *see* Riemann's Concordance Assumption; Riemannian space–Weyl space
 of time (sense of transported vectors), 798–796
 Pathological metrics. *See* Metric, pathological
 Pendulum model universe, 200, 204, 206
 Perceptual
 space. *see* Visual space
 threshold for time awareness, 633–637, 641–642, 880
 Perceptualism (Whiteheadian anti-bifurcationism), 11n, 48–65, 344–345, 425–427
 Periodicity objection (in statistical mechanics), 240–242, 244
 Perturbational influences. *See* Deforming influences
 Phase space (in statistical mechanics), 205, 237–239, 262, 267, 647–656, 660, 663
 Phenomenological level (in statistical mechanics). *See* Macro-state; Thermodynamics, classical; Partitioning of phase space
 Philosophy
 as conceptual clarification, 351–352, 709–711, 725–726, 755
 as purported extra-scientific cognition, 147–151
 See also Ordinary language philosophy
 Photons. *See* Bose-Einstein statistics; Theory of light, corpuscular; Light ray
 Physical
 possibility, 187, 203, 207, 826–827 (*see also* Causal connectibility)
 space (-time), *see* Space-time, physical vs. mathematical
 Pi, printing of decimal representation of π , 643–644. *See also* Infinity machines

- Pneumonia, as predictor of death, 297–298
- Point-set theory, 158–176, 529
- Points, 458–460, 463, 479–495, 499, 504, 522–523, 535–538, 812n. *See also* Elements of a manifold
- Polar coordinates. *See* Coordinates, polar
- Possibility. *See* Physical possibility
- Postulates. *See* Axioms
- Pragmatics, 131, 231–232, 326n., 837–838. *See also* Simplicity of theory
- Precession of Mercury's perihelion, 575, 576–579, 581–584
- Precognition, 287
- Prediction, 227, 232–234, 281–312, 320, 322, 832
epistemic failure of, in a probabilistic world, 321, 324
See also Retrodiction; Advance indicators; *H*-explanation and *H*-prediction
- Pre-geometry, Wheeler's conception of, 729–730, 757, 759–760. *See also* Relational theory of space-time
- Presentness. *See* 'Now'; Becoming
- Preservability of an hypothesis, 108–123, 139–144, 456, 587, 806–807. *See also* Duhemian thesis
- Presuppositional method. *See* Apriorism
- Primitive relations, 180, 191–193, 204–205, 230, 463, 465
- Probability, 589, 618–620, 650–659, 661–662, 664, 819
metric, in statistical mechanics, 238–243, 244n., 247, 262, 829–830
- Programmatic theses, as contrasted with elaborated ones, 110, 114–115, 337, 588, 608–610, 755, 776–777, 779, 826–827
- Progression. *See* Infinite series
- 'Progressive' metric of Massey. *See* Metric, 'pathological'
- Projective geometry and projective structure ($\mathcal{P}$), 121, 735n., 736, 739–744, 746–748, 751, 761, 764
- Properties:
general (purely qualitative), 505–506, 564, 845
language-independence of, 506
See also Relational properties;
Equivalence relation
- Psychoanalytic psychotherapy, 849
- Psychological:
vs. logical grounds, *see* Discovery,
context of vs. context of justification
space, *see* Visual space; Perceptualism;
Conscious organisms
time, 5n., 56–57, 59, 192–193, 221–232, 251, 263–264, 278, 281, 289–290, 314–319, 324–329 (*see also* Perceptual threshold for time awareness; 'Now'; Mind-dependence of becoming)
- Pulsars, 453
- Pythagorean infinitesimal metric behavior of rods, 732, 735, 742
- Quantized space. *See* Space, discrete
- Quantum mechanics, 236, 241n., 589, 729, 737, 751, 819, 836
metrogenic irreversibility in, 245–253, 298–299, 319–324, 831
providing a unit for phase space, 647
relation to relativity theory, xix, 502–503, 725, 789
space of, 336, 496, 503, 631, 739, 789
time-symmetry of, 240
- Quasi-closed systems. *See* Branch system
- Quasi-coincidence, 193n., 344
- Quasi-continuum, 17–18
- Quasi-disjoint intervals. *See* Intervals, quasi-disjoint
- Quasi-ergodic hypothesis, 238, 257, 262, 267, 662, 829–830
epistemic construal of denied, 238n.
- Quasi-Newtonian world, 370, 670, 684, 689–693, 695–696, 702, 727n., 834
- R. *See* Riemann's Concordance Assumption
- Radar ranging. *See* Light rays as means of signal synchrony
- Random occurrence of micro-states:
in a space-ensemble of macro-states, 256–263, 280, 656–657, 665n., 798
in a time-ensemble of macro-states, *see* Quasi-ergodic hypothesis
- Rational numbers. *See* Numbers, rational

- Rational and empirical factors in physical theory:
re axiomatization of theory, 122–123
re genesis and justification of theory, 382, 835
- Rationalism. *See* Apriorism
- Rationality. *See* Methodological justifiability
- RCA. *See* Riemann's Concordance Assumption
- RCS. *See* Synchrony, rule of committed
- Realism vs. anthropomorphism, 646–648
See also Entropy, anthropomorphic vs. realist construal of
- Records. *See* Traces
- Reciprocity between reference frames, principle of, 681–683
- Rectangular coordinates. *See* Coordinates, Cartesian
- Recurrence, 197–200, 202, 206. *See also* Identity of indiscernibles
- Refutation of hypotheses, 106–108, 111, 113. *See also* Falsifiability; Evidence, disconfirmatory
- Reichenbach's *c* notation. *See* Synchrony, standard and non-standard
- Relational properties, 63, 214–215, 404–405n., 451, 507, 512, 520, 522, 563–566, 844
 dyadicity vs. externality of, 505–507, 845–848
 intensional construal of, 451, 505–507, 509, 512, 513, 524, 550, 564–565
- Relational theory of space-time, 4–9, 42–43, 345–346, 352, 404–405, 408, 418–422, 729, 823, 841. *See also* Causal theory of time
- Relativity:
 of geometry, according to Reichenbach, 98–105
 principle of, 371, 374, 670, 678–683, 711, 725
 of simultaneity, 30, 204, 327, 343, 352–353, 359–368, 370–372, 377, 396, 401, 666, 693, 695, 700, 712–714, 834
- Relativity theory (TR), xvii–xx, xxii, 105, 179, 646
 axiomatization of, 192, 416, 714, 734
 continuous space of, 482–501, 505, 631, 753
 failure to incorporate all branches of physics, xix, 788–789
 general (GTR), relation to special (STR), 408, 419, 420, 843–844, 846, 789, 806, 841–842
 general, clock transport synchrony in, 669, 671, 707, 851
 general, dynamically-based metric of, 9, 55–56, 92–93, 123, 418–424, 425–428, 467, 501–505, 535–539, 545–547, 728, 783–764, 787, 840, 842–843
 general, experimental test of, 572, 574–584, 586
 general, space of, *see* Riemannian space
 general, spatial congruence in, 25, 92–93, 121, 123, 147, 473, 504, 551, 560, 562–563, 612, 730–751
 general, topology of time in, 201, 788–792, 798–800, 822, 824, 827, 831
 rotating disk in, 31n., 78–79, 542–545, 562, 604–606, 671, 673n., 707, 806
 special (STR), 71–72, 341–417, 453, 593, 598–601, 603, 606
 special, history of, 341, 342, 343, 369–383, 387, 393–394, 399, 408–409, 709–727, 834–839
 special, pedagogy of, 711–712, 714–715, 726
 special, simultaneity in, 29, 35, 204, 315n., 334, 341, 342–370, 425, 451, 561, 666–671, 675–707, 712–715, 834, 840, 851–854
 special, space of, *see* Space-time, Minkowskian
 special, topology of time in, 192, 203, 204, 317–318, 326–327, 824–827
 special, velocity addition law in, 369n., 375–377, 715
 temporal congruence in, 22, 23, 34–35, 66, 77–80, 454–455
 time-symmetry of, 216, 218
See also Newtonian mechanics, relation to relativity theory
- Relevance, semantic and non-semantic, 589, 594–607, 611
- Retrocausality, xxiii, 189, 825–827

- Retrodiction, 233-236, 243-244, 281-296, 299, 302, 312, 829
 epistemic asymmetry with prediction, 284, 295-296, 306
See also Prediction; Traces
- Reversibility, 180, 185-190, 193n., 206-208, 218, 233, 237, 239-240, 270-271
 objection to asymmetry of entropy statistics, 240-242, 244-245, 248
 strong (nomological), 189, 212-214
 weak (de facto reversedness), 189, 190, 212-214, 216, 217, 791
See also Time-symmetry; Irreversibility
- Riemann (-Christoffel) curvature tensor:
 first kind, *see* Curvature tensor, covariant
 second kind, *see* Curvature tensor, mixed
- Riemann's Concordance Assumption (RCA, or R), 222n., 503-504, 551(def.), 553, 555, 611-615, 618-620, 621n., 624-625, 696, 705, 743, 762-764, 795
- Riemann's Metrical Hypothesis (RMH), 8-16, 129n., 334, 455, 498(def.)-504, 512, 527(generalized statement), 526-531, 534, 540, 545-547, 548-550, 554-555, 559, 561, 562-563, 690, 751-759, 764, 786-788
 connection with dynamic conception of metrical field, *see* Metrical field, connection with RMH
 inductive support for, 498-501, 529-531, 753-755
- Riemannian:
 curvature, *see* Curvature, Riemannian metric (*R*-metric) and Riemannian space, 19, 33, 123, 157, 468, 469, 471(def.)-474, 476, 483, 489, 492, 507, 527-528, 531, 551, 559, 612, 732-751, 753, 755, 757, 773-775, 781-785, 793
- Rigidity. *See* Rods
- R*-metric. *See* Riemannian metric
- RMH. *See* Riemann's Metrical Hypothesis
- Robertson-Walker line element, 537-542
- Rods, rigid, as congruence standard, 11-15, 18, 21, 41-45, 54n., 56, 60-64, 82, 87, 90, 96-105, 109-110, 118, 119, 121, 123-125, 128-129n., 130, 132-146, 151, 156, 411-417, 452-453, 455, 474, 540-546, 549-553, 555, 556, 558, 559, 602-607, 611-624, 742, 786, 844, 848
 and clocks, isochronous, 3-4, 11-12, 26, 41, 52, 63, 78, 410, 416, 730-735, 740, 747-751, 753, 756, 762-764
- Rotating disk. *See* Relativity theory, rotating disk in
- Rule:
 of committed synchronism (RCS), *see* Synchronism, rule of committed
 of correspondence, *see* Coordinative 'definition'
- Scale factor, 88, 513, 521, 739, 777-778, 780, 785
- Schwarzschild:
 procedure, 119, 806
 solution or field, 420, 840
- Secular acceleration. *See* Acceleration, secular
- Self-congruence under transport, 3, 11, 13, 62, 78, 94, 104, 415, 452-455, 539, 541, 545, 558, 696, 705. *See also* Congruence standard; Conventionality of congruence
- Self-intersecting (Figure 8) structure:
 spatial, 331
 temporal, 203
- Semantics. *See* Interpretations of formal calculi; Meaning; Model theory; Conventionalism, trivial semantic
- Semantic stability, as necessary condition for non-triviality of Duhemian thesis, 111, 140, 143, 590, 591-594, 597-607
- Sentient observers. *See* Conscious organisms; Mind-dependence of becoming
- Separation:
 closure, *see* Betweenness, separation-closure
 null, *see* Metric, definite and indefinite;
 Light rays as null geodesics
 of space-time points, 472-473, 538, 542-544 (*see also* Distance)
 space-like, *see* Simultaneity, topological
- Serial ordering, 214-217, 218-219,

- 460, 463, 465, 468, 564, 566, 655, 797, 799
- S*-extrinsic (-intrinsic) metric. *See* Metric, extrinsic (intrinsic) to a manifold
- Signal (influence) chains. *See* Causal chains; Light rays as means for signal synchrony
- Similar figures, 102, 103, 157, 330, 331
- Simplicity of theory
- descriptive (convenience), 21, 23, 64, 66, 68, 72, 74, 77, 78, 79, 121, 122, 126, 133, 157, 356, 360, 367, 454, 556, 667
 - inductive, 151, 335, 355, 356, 716, 807
- Simultaneity:
- absolute, *see* Newtonian mechanics
 - indefinitely fast causal chain, in: Quasi-Newtonian world
 - metrical, 29, 32, 35, 40, 195, 204, 209, 223, 230, 315n., 327, 334, 341, 342, 368, 369, 372, 377, 410, 411, 666, 707, 712, 714, 834, 851
 - topological, 28, 31, 187, 188, 191, 193, 195, 203(def.), 204, 206, 207, 208, 319, 348, 353, 356, 366, 683, 689, 706, 733, 791, 792, 823
- See also* Conventionality of simultaneity; Synchronization; Clock; Relativity of simultaneity; Relativity theory, special, simultaneity in
- Singletons. *See* Points; Elements; Instants; Events
- Solar oblateness, 572, 578, 584, 586, 629n
- Solids, 128, 129n
- Sophist doctrine, 231n
- Space(-time):
- bounded by extreme elements (*E*-space and *F*-space), 461, 463, 465, 466, 479, 484, 493, 495, 508, 518, 522, 523, 525, 535, 554, 555, 564, 565
 - continuous, 8, 11, 12n., 16, 18, 41, 62, 91, 93, 94, 119, 129n., 144, 158, 176, 334, 337, 411, 426, 452, 454, 455, 458, 460, 461, 463(def.), 465, 469, 475, 476, 478, 480, 481, 482, 484, 485, 488, 489, 495, 504, 526, 534, 545, 553, 554, 566, 631, 636, 690, 751, 753, 786, 788, 812, 820, 843, 846
 - curved, *see* Curvature of space
 - denumerable dense, 10, 175, 335, 337, 460, 463(def.), 468, 469, 476, 481, 482, 484, 486, 488, 489, 517, 519, 531, 538, 559, 509, 814, 816, 820, 846
 - discrete (granular), 8, 9, 48, 208, 336, 486, 488, 460, 463, 467(def.), 465, 466, 470, 476, 477n., 479, 498, 500, 503, 508, 518, 526, 538, 840, 848, 848, 550, 556, 558, 559, 636n., 810, 843, 846
 - infinitude of, 244, 262, 267, 271, 275, 277, 420, 423, 424
 - Minkowskian, 30, 318, 326, 328, 464, 472, 473, 492, 839n., 842, 843, 844, 846, 847, 789, 806, 821, 825, 841, 842
 - physical vs. mathematical, 173, 174, 459, 460, 462, 468, 466, 467, 478, 495, 497, 506, 508, 509, 511, 512, 517, 518, 566, 782, 810, 820
 - real geometry of, 147, 151, 760
 - Riemannian, *see* Riemannian metric
 - spaces distinguished from manifolds, 459(def.), 485, 486
 - space distinguished from space-time, 328, 535, 541, 543, 544
- See also* Congruence; Geometry; Metric; Topology
- Spacelike separation. *See* Simultaneity, topological
- Space-theory of matter. *See* Geometrodynamics (GMD)
- Special theory of relativity (STR). *See* Relativity theory, special
- Spherical geometry. *See* Geometry, spherical
- Spontaneous
- concatenation of initial conditions, 265, 271
 - generation (SG), in biology, 571, 574
- Staccato run. *See* Z-run, staccato
- Standard meter, 36, 454, 455, 556, 563
- Statistical mechanics, 192, 219, 220, 222, 224n., 230n., 235, 264, 266, 320, 646, 665, 754, 790, 791, 798, 799, 829, *See also* Probability; Micro-state; Macro-state; Entropy; Phase space; Quantum mechanics
- Steady-state. *See* Cosmology, steady-state
- Stellar parallax, 24, 109, 110, 115, 116, 119, 121, 124, 126, 149, 806, 808

- STR. *See* Relativity theory, special
- Straight line, 13, 20, 23, 32, 33, 81, 102, 115–118, 120, 135. *See also* Geodesic
- Substance. *See* Ontology
- Super-denumerable. *See* Number, real:
Space, continuous; Order, dense
continuous
- Superluminal causal chains. *See* Tachyons
- Superposition objection, as defense of
Duhemian thesis, 616, 618–620, 621,
622–623n.
- Supertasks. *See* Infinity machines
- Symbol-using organisms, 284–286, 287
- Synchronization (synchronism) of clocks,
29–31, 344–345, 346–347, 355–365,
411–414, 666–708
rule of committed synchronism (RCS),
and generalized rule of committed
synchronism (GRCS), 672–678,
852–854
standard ($\epsilon = \frac{1}{2}$) and non-standard ($\epsilon \neq \frac{1}{2}$),
30, 353–367, 396–397, 666–683,
693–707, 743, 840, 852–853
See also Simultaneity
- τ vs. t time scales. *See* Clock, atomic vs.
astronomical
- T vs. t time scales. *See* Clock, diurnal vs.
ephemeris
- Tachyons (superluminal causal chains),
xxiii, 350, 687, 822, 824–827
- Tangent vector, 794–796
- Teleology. *See* Mechanism vs. teleology
- Telling the spatial or temporal story, 476,
481, 483–495, 515, 555–556
- Temperature. *See* Thermometrics
- Temporal:
counter-directedness of systems,
224–230, 260, 261, 280
separation, *antonym of* simultaneity, *q.v.*
See also Time
- Tensed verbs, 184–185. *See also* Becoming
- Tensors, 769, 782, 784
energy-momentum, 546–547, 842
See also Curvature tensor; Metric
tensor
- Ternary set (Cantor discontinuum), 164
- Testability, 81–82, 96, 125, 387–388, 392,
550–553, 556, 717–721, 837–839. *See*
also Falsifiability; Verifiability;
Empiricism; *Ad hoc*
- Theology, 230n.
- Theorema egregium* of Gauss's theory of
surfaces, 781
- Theoretical:
language, 111, 591, 593, 596–597
system, 106, 108, 111, 115, 124, 133,
135–139, 284–289, 387–388, 576,
585–586, 590, 807–808 (*see also*
Duhemian thesis)
terms, 134, 387–388, 451, 717 (*see also*
Interpretations of formal calculi)
- Theory-laden findings, 136, 383, 387, 622
- Thermal expansion. *See* Deforming in-
fluences
- Thermodynamics, second law of, 186n.,
187
classical (phenomenological), 216, 217,
219–236, 251n., 372
statistical. *see* Statistical mechanics
See also Entropy; Irreversibility
- Thermometrics, 132, 562–563, 681,
691–692, 810–811
- Three-dimensionality of space, 330–334,
820, 833–834
- Time:
anisotropy of. *see* Anisotropy of time
betweenness for. *see* Betweenness,
temporal; Causal theory of time
'direction' of. *see* Anisotropy of time;
Becoming
discrete, 208, 636–637
'flow' of. *see* Becoming
metric for. *see* Chronometry; Clock;
Congruence class for time intervals;
Conventionality of congruence,
temporal; Newtonian mechanics,
time-metric of Simultaneity, metrical
openness of. *see* Closed time; Open time
See also Temporal counterdirectedness
of systems; Space-time
- Time-like geodesics. *See* Freely falling
particles, massive
- Time-orientability and time-orientation
of space-time, 788–800, 831
- Time-orthogonality. *See* Coordinates,
time-orthogonal
- Time-symmetry, 214, 216, 224, 240, 243,

- 244, 263, 272, 273, 277, 280, 300, 305,
 309, 312. *See also* Isotropy of time.
 Reversibility
 'Topological' condition for light ($0 < \epsilon < 1$).
 353, 354, 360, 671, 672, 676, 677, 703,
 706, 707
 Topology:
 senses of the term, figure-to-figure vs
 automorphic, 529, 530
 senses of the term, non-metrical vs
 invariant under homeomorphism,
 820, 821, 828
 of space, 96, 97, 104, 161, 201, 330, 337,
 497, 499, 529, 535, 536, 820, 833
 of time, *see* Causal theory of time.
 Anisotropy of time; Becoming,
 Betweenness, temporal
 of time, macro-character of, 207
 TR. *See* Relativity theory
 Traces (records of past interactions,
 veridical retrodictive indicators), 243,
 263, 264, 282, 289, 295, 297, 300, 312,
 320, 322, 325
 Transitivity:
 of inductive support, 624
 of synchrony, 673, 678, 852, 853
 Transported congruence standard. *See*
 Congruence standard, Self-congru-
 ence; Rods; Clock; Riemann's Con-
 cordance Assumption
 Triangle, 102, 103, 109, 110, 116
 inequality for distance functions, 468,
 471, 473, 517
 Triviality. *See* Conventionality, trivial.
 Duhemian thesis, charge of triviality
 against; Metric, nontriviality of;
 Metrizability, alternative, nontrivial
 construal of
 Trivial Semantic Conventionism (TSC).
 See Conventionality, trivial seman-
 tic
 Unified field theory, 787
 Unique metric purportedly needed for
 curvature. *See* Curvature, thesis of
 Requirement of a Unique Metric
 (CRUM)
 'Universal forces' of Reichenbach, 82, 97,
 101, 105, 118, 120, 703, 704
 Universe. *See* Cosmology; Cosmogony;
 Space-time, infinitude of; Entropy
 of the universe; Expansion of the
 universe; Model universes
 Unreversedness. *See* Irreversibility, de
 facto
 Unspecifiability of clues, according to
 Polanyi, 378, 379, 382, 384, 385, 834, 837
 Use vs. mention, 37

 Vagueness, 131
 Van der Waals's law, 34, 38, 41
 Variations, calculus of, 13, 14n, 20, 23
 Verification of hypotheses, 106, 108, 110,
 111, 133, 136, 137, 326n, 569, 570,
 572n, 586, 587, 624, 625, 718, 722,
 819, 849. *See also* Testability,
 Falsifiability
 Visual space, 6, 8, 152, 157, 174
 Volition (choice, decision), 647, 648, 656,
 663, 825, 826

 Wave equation, 332, 333, 725
 of Schrodinger, 233, 247, 250, 298
 Weyl space, 741, 744
 B^* -manifold, 527(def.) 530
 World lines, 201, 537, 541, 545, 824. *See*
 also Causal chains, genidentical

 Zeno's paradoxes
 of metrical extension, 158, 176, 335,
 532, 533, 808, 820
 of motion, 57n, 159, 180, 208,
 630, 645, 808, 828, 849, 851
 Z-interval, Z-sequence, and Z-stop,
 632, 645
 Z-run:
 legato, 632, 643, 850
 marking in the course of, 638, 639,
 642, 643
 staccato, 632, 638, 644, 850

SYNTHESE HISTORICAL LIBRARY

Texts and Studies
in the History of Logic and Philosophy

Editors:

N. KRETZMANN (Cornell University)
G. NUCHELMANS (University of Leyden)
L. M. DE RIJK (University of Leyden)

LEWIS WHITE BECK (ed.), *Proceedings of the Third International Kant Congress*. 1972,
XI + 718 pp.

‡KARL WOLF and PAUL WEINGARTNER (eds.), *Ernst Mally: Logische Schriften*. 1971,
X + 340 pp.

‡LEROY E. LOEMKER (ed.), *Gottfried Wilhelm Leibnitz: Philosophical Papers and Letters*.
A Selection Translated and Edited, with an Introduction. 1969, XII + 736 pp.

‡M. T. BEONIO-BROCCHIERI FUMAGALLI, *The Logic of Abelard*. Translated from the
Italian. 1969, IX + 101 pp.

Monographs on Epistemology, Logic, Methodology,
Philosophy of Science, Sociology of Science and of Knowledge, and on the
Mathematical Methods of Social and Behavioral Sciences

Editors

DONALD DAVIDSON (The Rockefeller University and Princeton University)

JAAKKO HINTIKKA (Academy of Finland and Stanford University)

GABRIEL NUCHTEMANN (University of Leyden)

WESLEY C. SALMON (Indiana University)

ILKKA NIINILUOTO and RAIMO TUOMELA: *Theoretical Concepts and Hypothetico-Inductive Inference*. 1973, VIII + 264 pp.

ROLAND FRAISSÉ: *Course of Mathematical Logic*. Volume I: *Relation and Logical Formula*. 1973, XVI + 186 pp. Also available as a paperback.

RADU J. BOGDAN and ILKKA NIINILUOTO (eds.): *Logic, Language and Probability*. A selection of papers contributed to Sections IV, VI and XI of the Fourth International Congress for Logic, Methodology, and Philosophy of Science, Bucharest, September 1971. 1973, X + 323 pp.

GLENN PEARCE and PATRICK MAYNARD (eds.): *Conceptual Change*. 1973, XII + 282 pp.

MARIO BUNGE (ed.): *Exact Philosophy: Problems, Tools, and Goals*. 1973, X + 214 pp.

ROBERT S. COHNS and MARX W. WARTOFSKY (eds.): *Boston Studies in the Philosophy of Science*, Volume IX: *A. A. Zinov'ev: Foundations of the Logical Theory of Scientific Knowledge (Complex Logic)*. Revised and Enlarged English Edition with an Appendix by G. A. Smirnov, E. A. Sidorenka, A. M. Fedina, and L. A. Bobrova. 1973, XXII + 301 pp. Also available as a paperback.

K. J. J. HINTIKKA, J. M. E. MORAVCSIK, and P. SUPPES (eds.): *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*. 1973, VIII + 526 pp. Also available as a paperback.

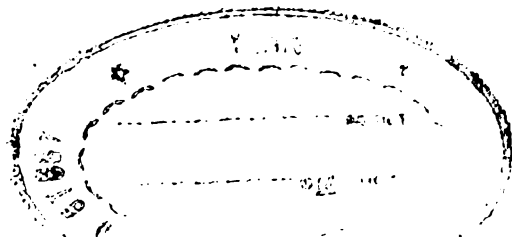
WILLARD V. HUMPHREYS, JR. (ed.): *Norwood Russell Hanson: Constellations and Conjectures*. 1973, X + 282 pp.

MARIO BUNGE: *Method, Model and Matter*. 1973, VII + 196 pp.

MARIO BUNGE: *Philosophy of Physics*. 1973, IX + 248 pp.

LADISLAV TONDÉ: *Boston Studies in the Philosophy of Science*, Volume X: *Scientific Procedures*. 1973, XIII + 268 pp. Also available as a paperback.

SÖREN STENLUND: *Combinators, λ -Terms and Proof Theory*. 1972, 184 pp.



- DONALD DAVIDSON and GILBERT HARMAN (eds.), *Semantics of Natural Language*. 1972, X + 769 pp. Also available as a paperback.
- MARTIN STRAUSS, *Modern Physics and Its Philosophy. Selected Papers in the Logic, History, and Philosophy of Science*. 1972, X + 297 pp.
- ‡STEPHEN TOULMIN and HARRY WOOLF (eds.), *Norwood Russell Hanson: What I Do Not Believe, and Other Essays*. 1971, XII + 390 pp.
- ‡ROBERT S. COHEN and MARX W. WARTOFSKY (eds.), *Boston Studies in the Philosophy of Science*. Volume VIII: *PSA 1970. In Memory of Rudolf Carnap* (ed. by Roger C. Buck and Robert S. Cohen). 1971, LXVI + 615 pp. Also available as a paperback.
- ‡YEHOŠUA BAR-HILLEL (ed.), *Pragmatics of Natural Languages*. 1971, VII + 231 pp.
- ‡ROBERT S. COHEN and MARX W. WARTOFSKY (eds.), *Boston Studies in the Philosophy of Science*. Volume VII: *Milič Čapek: Bergson and Modern Physics*. 1971, XV + 414 pp.
- ‡CARL R. KORDIG, *The Justification of Scientific Change*. 1971, XIV + 119 pp.
- ‡JOSEPH D. SNEED, *The Logical Structure of Mathematical Physics*. 1971, XV + 311 pp.
- ‡JEAN-LOUIS KRIVINE, *Introduction to Axiomatic Set Theory*. 1971, VII + 98 pp.
- ‡RISTO HILPINEN (ed.), *Deontic Logic: Introductory and Systematic Readings*. 1971, VII + 182 pp.
- ‡EVERET W. BETH, *Aspects of Modern Logic*. 1970, XI + 176 pp.
- ‡PAUL WEINGARTNER and GERHARD ZECHA, (eds.), *Induction, Physics, and Ethics. Proceedings and Discussions of the 1968 Salzburg Colloquium in the Philosophy of Science*. 1970, X + 382 pp.
- ‡ROLF A. EBERLE, *Nominalistic Systems*. 1970, IX + 217 pp.
- ‡JAAKKO HINTIKKA and PATRICK SUPPES, *Information and Inference*. 1970, X + 336 pp.
- ‡KAREL LAMBERT, *Philosophical Problems in Logic. Some Recent Developments*. 1970, VII + 176 pp.
- ‡P. V. TAVANEC (ed.), *Problems of the Logic of Scientific Knowledge*. 1969, XII + 429 pp.
- ‡ROBERT S. COHEN and RAYMOND J. SEEGER (eds.), *Boston Studies in the Philosophy of Science*. Volume VI: *Ernst Mach: Physicist and Philosopher*. 1970, VIII + 295 pp.
- ‡MARSHALL SWAIN (ed.), *Induction, Acceptance, and Rational Belief*. 1970, VII + 232 pp.
- ‡NICHOLAS RESCHER *et al.* (eds.), *Essay in Honor of Carl G. Hempel. A Tribute on the Occasion of his Sixty-Fifth Birthday*. 1969, VII + 272 pp.
- ‡PATRICK SUPPES, *Studies in the Methodology and Foundations of Science. Selected Papers from 1911 to 1969*. 1969, XII + 473 pp.
- ‡JAAKKO HINTIKKA, *Models for Modalities. Selected Essays*. 1969, IX + 220 pp.
- ‡D. DAVIDSON and J. HINTIKKA (eds.), *Words and Objections: Essays on the Work of W. V. Quine*. 1969, VIII + 366 pp.
- ‡J. W. DAVIS, D. J. HOCKNEY and W. K. WILSON (eds.), *Philosophical Logic*. 1969, VIII + 277 pp.
- ‡ROBERT S. COHEN and MARX W. WARTOFSKY (eds.), *Boston Studies in the Philosophy of Science*. Volume V: *Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968*. 1969, VIII + 482 pp.

- ‡ROBERT S. COHEN and MARX W. WARTOFSKY (eds.), *Boston Studies in the Philosophy of Science* Volume IV *Proceedings of the Boston Colloquium for the Philosophy of Science 1966-1968*, 1969, VIII + 537 pp.
- ‡NICHOLAS RESCHER, *Topics in Philosophical Logic*, 1968, XIV + 347 pp.
- ‡GÜNTHER PATZIG, *Aristotle's Theory of the Syllogism: A Logical-Philological Study of Book A of the Prior Analytics*, 1968, XVII + 215 pp.
- ‡C. D. BROAD, *Induction, Probability, and Causation: Selected Papers*, 1968, XI + 296 pp.
- ‡ROBERT S. COHEN and MARX W. WARTOFSKY (eds.), *Boston Studies in the Philosophy of Science* Volume III *Proceedings of the Boston Colloquium for the Philosophy of Science 1964-1966-1967*, XIIX + 489 pp.
- ‡GUIDO KUNG, *Ontology and the Logical Analysis of Language: An Enquiry into the Contemporary Views on Universals*, 1967, XI + 210 pp.
- *EVERT W. BETH and JEAN PIAGET, *Mathematical Epistemology and Psychology*, 1966, XXII + 326 pp.
- *EVERT W. BETH, *Mathematical Thought: An Introduction to the Philosophy of Mathematics*, 1965, XII + 208 pp.
- ‡PAUL LORENZEN, *Formal Logic*, 1965, VIII + 123 pp.
- ‡GEORGES GURVITCH, *The Spectrum of Social Time*, 1964, XXVI + 152 pp.
- ‡A. A. ZINOV'EV, *Philosophical Problems of Many-Valued Logic*, 1963, XIV + 155 pp.
- ‡MARX W. WARTOFSKY (ed.), *Boston Studies in the Philosophy of Science* Volume I: *Proceedings of the Boston Colloquium for the Philosophy of Science, 1961-1962, 1963*, VIII + 212 pp.
- ‡B. H. KAZEMIER and D. VUYSJE (eds.), *Logic and Language: Studies dedicated to Professor Rudolf Carnap on the Occasion of his Seventieth Birthday*, 1962, VI + 256 pp.
- *EVERT W. BETH, *Formal Methods: An Introduction to Symbolic Logic and to the Study of Effective Operations in Arithmetic and Logic*, 1962, XIV + 170 pp.
- *HANS FREUDENTHAL (ed.), *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences. Proceedings of a Colloquium held at Utrecht, The Netherlands, January 1960, 1961*, VI + 194 pp.
- ‡P. L. GUIRAUD, *Problèmes et méthodes de la statistique linguistique*, 1960, VI + 146 pp.
- *J. M. BOCHENSKI, *A Precise of Mathematical Logic*, 1959, X + 100 pp.

Sole Distributors in the U.S.A. and Canada:

*GORDON & BREACH, INC., 440 Park Avenue South, New York, N.Y. 10016

‡HUMANITIES PRESS, INC., 303 Park Avenue South, New York, N.Y. 10010

