VAN NOSTRAND REINHOLD MATHEMATICAL STUDIES #30

Notes on OPERATOR THEORY

Peter A. Fillmore
Indiana University

| MATHEMATICAL STUDIES

ors

ndiana University

e University of Michigan

; ON BOOLEAN ALGEBRAS

PTIC BOUNDARY VALUE PROBLEMS

DIFFERENTIAL GEOMETRY

POLOGY AND ORDER

ES ON SPECTRAL THEORY

NOTES ON LOGIC

; ON CHOQUET'S THEOREM

SES IN SET THEORY

...RY OF FUNCTIONS OF A COMPLEX VARIABLE

Lars V. Ahlfors—**LECTURES ON QUASICONFORMAL MAPPINGS**

J. Peter May—**SIMPLICIAL OBJECTS IN ALGEBRAIC TOPOLOGY**

W. H. J. Fuchs—**TOPICS IN THE THEORY OF FUNCTIONS OF ONE COMPLEX VARIABLE**

Lennart Carleson—**SELECTED PROBLEMS ON EXCEPTIONAL SETS**

Leopoldo Nachbin—**ELEMENTS OF APPROXIMATION THEORY**

S. S. Chern—**COMPLEX MANIFOLDS WITHOUT POTENTIAL THEORY**

F. P. Greenleaf—**INVARIANT MEANS ON TOPOLOGICAL GROUPS**

H. Widom—**LECTURES ON INTEGRAL EQUATIONS**

A. K. Aziz, Gen. Ed.—**LECTURE SERIES IN DIFFERENTIAL EQUATIONS, Vol. I**

A. K. Aziz, Gen. Ed.—**LECTURE SERIES IN DIFFERENTIAL EQUATIONS, Vol. II**

Harold Widom—**LECTURES ON MEASURE AND INTEGRATION**

Shaul R. Foguel—**THE ERGODIC THEORY OF MARKOV PROCESSES**

Makoto Ohtsuka—**DIRICHLET PROBLEM, EXTREMAL LENGTH AND PRIME ENDS**

Hans Samelson—**NOTES ON LIE ALGEBRAS**

Harold S. Shapiro—**SMOOTHING AND APPROXIMATION OF FUNCTIONS**

Robert Osserman—**A SURVEY OF MINIMAL SURFACES**

Joseph J. Rotman—**NOTES ON HOMOLOGICAL ALGEBRA**

Leslie C. Glaser—**GEOMETRICAL COMBINATORIAL TOPOLOGY, Vol. I**

Leslie C. Glaser—**GEOMETRICAL COMBINATORIAL TOPOLOGY, Vol. II**

Nathaniel A. Friedman—**INTRODUCTION TO ERGODIC THEORY**

Peter A. Fillmore—**NOTES ON OPERATOR THEORY**

J. C. Abbott, Gen. Ed.—**TRENDS IN LATTICE THEORY**

Notes on OPERATOR THEORY

by

PETER A. FILLMORE
Indiana University

VAN NOSTRAND REINHOLD COMPANY
NEW YORK CINCINNATI TORONTO LONDON MELBOURNE

Van Nostrand Reinhold Company Regional Offices:
Cincinnati, New York, Chicago, Millbrae, Dallas

Van Nostrand Reinhold Company Foreign Offices:
London, Toronto, Melbourne

Manufactured in the United States of America

Published by Van Nostrand Reinhold Company
450 West 33rd Street, New York, N.Y. 10001

Published simultaneously in Canada by
D. Van Nostrand Company (Canada), Ltd.

10 9 8 7 6 5 4 3 2 1

PREFACE

These notes represent the substance of a course of lectures I delivered at Indiana University in 1967. The audience was familiar with the basic theory of Hilbert spaces and operators up through the spectral theorem and the theory of spectral multiplicity. For many interesting and helpful conversations I am indebted to R. G. Douglas, and to my colleagues Arlen Brown, J. G. Stampfli, D. M. Topping and J. P. Williams.

P.A.F.

Bloomington, Indiana
November, 1968

CONTENTS

INTRODUCTION

Many results concerning bounded linear operators in Hilbert space may be regarded as contributions to a structure theory for these operators. While it is unlikely that a theory as comprehensive as that for the finite-dimensional case exists in general, there is no doubt that substantial progress is possible. Evidence that comes easily to mind includes (in addition to the spectral theorem) Dunford's theory of spectral operators and the successful analysis by numerous workers of operators that are close (in various senses) to being normal.

One can identify several themes in this work. A natural one is the attempt to push finite-dimensional results to the infinite-dimensional situation. A well-known example of this is the spectral theorem, which results from the diagonal form for self-adjoint matrices when finite-dimensionality is dropped. Again, an extensive theory initiated by Livsic in 1954 is concerned with results inspired by the subdiagonal form for arbitrary matrices. Here one comes upon a major obstacle: the existence of a subdiagonal form for a matrix amounts to the existence of a chain of invariant subspaces, one of each dimension, and in general it is not known whether invariant subspaces must exist. Indeed, some authors are inclined to doubt it. Even granting a good supply of invariant subspaces, what to do with them is in general far from clear. Of course, suitable discrete chains of invariant subspaces will lead (as in the finite-dimensional case) to a representation of the operator as an infinite subdiagonal matrix. Unfortunately such chains need not exist, even for self adjoint operators (because of the possible dearth of eigenvectors). Thus even when blessed with a profusion of invariant subspaces, one must be prepared to contend with dense chains of them.

To obtain the necessary invariant subspaces, Livsic assumes that the imaginary part of the operator belongs to the trace class. This also implies that the non-real spectrum consists of eigenvalues clustering only on the real axis. This discrete part of the operator may be treated by methods that are essentially finite-dimensional, leaving as the interesting case an operator with real spectrum. Here the reduction to subdiagonal form consists in representing the operator, on an L^2-space of vector-valued functions on an interval $[0,\ell]$, as the sum of a multiplication by a non-decreasing function (the self adjoint part) and a Volterra integration operator with an operator-valued kernel (the quasinilpotent part).

Another important and fruitful theme, developed by Sz.-Nagy and many others, relates the structure of a contrative operator T to the behavior of the sequence $I, T, T^2, \ldots$ of its iterates. Perhaps the major actor in this approach is the shift operator, an operator that often seems the embodiment of the typically infinite-dimensional. One need not look far to explain its appearance here: if each vector x is made to correspond to the sequence $(x, Tx, T^2x, \ldots)$, then Tx corresponds to $(Tx, T^2x, \ldots)$, which is the previous sequence shifted one place to the left. Suitable modifications of this lead easily to the following model for contractions: every contraction is unitarily equivalent to an operator obtained by restricting a coisometry (the adjoint of an isometry) to an invariant subspace. Unhappily this elegant result is often of limited utility, a fact that arises from the bewildering variety of invariant subspaces possessed by a coisometry. The situation is rather frustrating and even ludicrous; with this model the invariant subspace problem mentioned above obviously can be settled by determining whether the minimal non-zero invariant subspaces of a coisometry must be one-dimensional. In spite of the apparently simple nature of coisometries, even relatively crude information of this sort is available only when the nullity of the coisometry is finite. Before leaving this topic it should be mentioned that whereas the Livisic theory attempts to view an operator as a perturbation of a self-adjoint operator, here it is regarded as a perturbation of an isometry. Thus a contraction T such that $I - T^*T$ is not too big has invariant subspaces.

This book is intended to be an introduction to structure theory for operators, and consequently attempts to provide a broad background. For proofs of a number of the deepest results the reader is referred to the literature. Although material is drawn from a wide variety of sources, the point of view is usually the second of those previously outlined. There are some new contributions, mostly of a modest nature, such as simplified proofs and improved points of view, as well as a few new results. From the foregoing discussion one might expect the stars of the book to be shift operators and invariant subspaces, and such is the case; on each page, if not explicitly present, they are not far behind the scenes.

We conclude this introductory section with a brief outline of the contents. After a section about hyponormal operators, which contains material not conveniently assembled elsewhere, the next two sections are devoted to the elementary facts about shift operators and their role in model theory. The basic theorems of Rota and of de Branges and Rovayak characterizing the parts of the backward shift up to similarity and up to unitary equivalence are obtained. With the relevance of invariant subspaces of shift operators thus established, this topic is taken up in the following two sections. Beurling's results relating the invariant subspaces of the simple shift and inner functions are derived and extended to shifts of countable multiplicity. This leads to the well-known formulation of the invariant subspace problem as a factorization problem for operator-valued functions.

In section 6 the Nagy-Foias model for contractions is developed. We obtain the coisometric extension in two ways: quickly, by displaying a matrix representation, and more laboriously, by means of a construction that reveals in greater detail the structure of the extension. From this follows the decomposition into unitary and completely nonunitary summands, as well as a related decomposition due to Foguel. Sections 7 and 8 contain numerous results concerning unitary dilations of contractions and contractive operator-valued mappings. The strong unitary dilation is obtained from the coisometric extension, the structure of this dilation is examined, and several applications are given. In section 8 we describe the beautiful circle of results, due to Bochner, Herglotz, Naimark, and Stone, dealing with positive-definite operator-valued functions, unitary representations, and positive operator-valued measures.

The ideas developed for the study of a single contraction (or rather, of its powers) can be applied to continuous one-parameter contractive semigroups, and this is the subject of section 9. The coisometric extension is constructed and used to deduce Cooper's Theorem on the structure of isometric semigroups. Then a quite distinct point of view is adopted, in which the semigroup is studied in terms of its infinitesimal generator and its cogenerator.

A number of results related through the concept of hyperinvariance are gathered in section 10. We show that the Volterra operator and certain weighted shifts are unicellular, and that the invariant subspaces of a unicellular operator are hyperinvariant. An operator that is quasi-similar to a unitary operator has nontrivial hyperinvariant subspaces, and from this several invariant subspace theorems for contractions are deduced. The concluding section is devoted to the recent proof of Arveson and Feldman that compact operators have nontrivial invariant subspaces.

SECTION 0

HYPONORMAL OPERATORS

Throughout these notes the term "operator" will be reserved for a bounded linear transformation on a Hilbert space, and a "subspace" will mean a closed linear manifold. Undoubtedly the best-understood operators are the normal operators. Because of this a recurrent theme in operator theory is the study of operators that in some sense are close to being normal. An interesting class of such operators arises from the following easily-stated but unsolved problem:

Problem: If N is a normal operator and $\mathfrak{M}$ an invariant subspace of dimension ≥ 2, is there an invariant subspace lying properly between $\{0\}$ and $\mathfrak{M}$?

Operators such as $N|\mathfrak{M}$, obtained by restricting a normal operator to an invariant subspace, are called *subnormal.* Thus the above problem is: does every subnormal operator have a nontrivial invariant subspace? This is a special case of one of the basic unsolved problems of operator theory:

Invariant Subspace Problem: Does every operator have a nontrivial invariant subspace?

Unlike the corresponding situation for self-adjoint operators, a subnormal operator need not be normal. For example, let ℓ^2 be the Hilbert space of doubly-infinite sequences $a = \{a_n\}$ of complex numbers such that $\|a\|^2 = \Sigma|a_n|^2 < \infty$, and let U be the bilateral shift: $(Ua)_n = a_{n-1}$. Then U is unitary, and the subspace $\mathfrak{M} = \{a | a_n = 0 \text{ for } n < 0\}$ is invariant, but $U|\mathfrak{M}$ is not normal. On the other hand, not every operator is subnormal. In fact, if $T = N|\mathfrak{M}$ is subnormal, then for all $x \in \mathfrak{M}$,

$$\|T^*x\| = \|P_{\mathfrak{M}}N^*x\| \leq \|N^*x\| = \|Nx\| = \|Tx\| ,$$

or equivalently, $TT^* \leq T^*T$. Such operators are called *hyponormal*; they constitute a class which contains properly the subnormals.

In some circumstances the problem mentioned above has an affirmative solution. The following theorem is due to Wermer [58].

I. If N is a normal operator with spectrum $\sigma(N)$ of area zero, and if $\mathfrak{M}$ is an invariant subspace of dimension ≥ 2, then $T = N|\mathfrak{M}$ has a nontrivial invariant subspace.

Proof: Let $\mathfrak{K}$ be the smallest subspace containing $\mathfrak{M}$ and reducing N. Since $N|\mathfrak{K}$ is a normal operator with spectrum contained in $\sigma(N)$, there is no harm in assuming that $\mathfrak{K}$ is the whole space. With this reduction it follows [24, Problem 157] that $\sigma(N) \subset \sigma(T)$. On the other hand, it is obvious that the point spectrum of N includes that of T, and that the that the same relation holds for continuous spectra. Since N has no residual spectrum, this implies that $R = \sigma(T) \setminus \sigma(N)$ lies in the residual spectrum of T. But it is easy to see that any operator with residual spectrum has a nontrivial invariant subspace, and so the case in which $\sigma(T) = \sigma(N)$ remains. The proof is completed by showing that in this situation T is actually normal; i.e., that $\mathfrak{M}$ reduces N. Since $\sigma(N)$ has zero area, the Hartogs-Rosenthal Theorem [27] implies that there are rational functions $r_n(z)$ converging uniformly to $\bar{z}$ on $\sigma(N)$, so that $r_n(N)$ converges to N^* in norm. This means that it will be sufficient to show that the resolvent R_λ of N leaves $\mathfrak{M}$ invariant. But if λ is in the resolvent set $\rho(N)$, then $\lambda \in \rho(T)$ and $(N-\lambda I)|\mathfrak{M} = T - \lambda I$, so that $(N-\lambda I)\mathfrak{M} = \mathfrak{M}$ and $R_\lambda \mathfrak{M} = \mathfrak{M}$.

This proof is due to J. G. Stampfli, as is the following observation.

II. If T is hyponormal, then $\{x \mid \|Tx\| = \|T\|\,\|x\|\}$ is an invariant subspace.

Proof: It can be supposed that $\|T\| = 1$. If $\|Tx\| = \|x\|$, then

$$\begin{aligned} \|T^*Tx - x\|^2 &= \|T^*Tx\|^2 - 2\mathrm{Re}(T^*Tx, x) + \|x\|^2 \\ &\leq \|Tx\|^2 - 2\|Tx\|^2 + \|x\|^2 \\ &\leq 0\,, \end{aligned}$$

and so $T^*Tx = x$. Since the converse is obvious, the above set is a subspace. Invariance results from the following computation: if $\|Tx\| = \|x\| = 1$, then

$$\|T^2x\| \geq \|T^*Tx\| \geq (T^*Tx, x) = \|Tx\|^2 = 1 \geq \|T^2x\|,$$

so that $\|T^2x\| = \|Tx\|$.

Many important properties of normal operators are valid for hyponormals. Some of these are proved below, and others may be found in the exercises.

III. Let T be hyponormal. Then

1. $T-\lambda I$ and T^{-1} are hyponormal,
2. $Tx=\lambda x$ implies $T^*x=\bar{\lambda}x$, and
3. $Tx=\lambda x$, $Ty=\mu y$, and $\lambda\neq\mu$ imply that x and y are orthogonal.

Proof: $(T-\lambda I)(T^*-\bar{\lambda}I) = TT^*-\lambda T^*-\bar{\lambda}T+|\lambda|^2 I \leq T^*T-\lambda T^*-\bar{\lambda}T+|\lambda|^2 I = (T^*-\bar{\lambda}I)(T-\lambda I)$, so that $T-\lambda I$ is hyponormal. If T is invertible and $TT^*\leq T^*T$, then

$$I\leq T^{-1}T^*TT^{*-1},\ T^*T^{-1}T^{*-1}T\leq I,$$

and

$$T^{-1}T^{*-1}\leq T^{*-1}T^{-1}.$$

The second statement is clear from the first, and for the third, $\lambda(x,y)=(Tx,y)=(x,T^*y)=\mu(x,y)$.

IV. If T is hyponormal then $\|T^n\|=\|T\|^n$ for $n=1,2,\ldots$, and consequently $r(T)=\|T\|$. (Here $r(T)$ denotes the spectral radius of T.)

Proof: The proof is by induction. First

$$\|T^n x\|^2=(T^*T^n x,\ T^{n-1}x)\leq\|T^*T^n x\|\ \|T^{n-1}x\|$$
$$\leq\|T^{n+1}x\|\ \|T^{n-1}x\|,$$

so that $\|T^n\|^2\leq\|T^{n+1}\|\ \|T^{n-1}\|$ for all $n\geq 1$. If $\|T^k\|=\|T\|^k$ for $1\leq k\leq n$, then

$$\|T\|^{2n}=\|T^n\|^2\leq\|T^{n+1}\|\ \|T^{n-1}\|=\|T^{n+1}\|\ \|T\|^{n-1},$$

and consequently $\|T\|^{n+1}\leq\|T^{n+1}\|$. The other inequality is true for any operator T. The second statement results from the fact that $r(T)$ is the limit of $\|T^n\|^{1/n}$ for any T [24, Prob. 74].

V. If T is hyponormal then $\|(T-\lambda I)^{-1}\| = 1/d$, where d is the distance from λ to the spectrum $\sigma(T)$ of T.

Proof: It is clear that $r((T-\lambda I)^{-1}) = 1/d$ for any operator T. But $(T-\lambda I)^{-1}$ is hyponormal by III, and so IV applies.

For the next result, recall that $\overline{W}(T)$ denotes the closure of the numerical range $\{(Tx, x) \mid \|x\| = 1\}$, and that $\overline{W}(T)$ is convex and contains $\sigma(T)$ [24, Problems 166 and 169].

VI. If T is hyponormal, then $\overline{W}(T)$ is equal to the convex hull of $\sigma(T)$.

Proof: By the above remarks we need only show that the convex hull of $\sigma(T)$ contains $\overline{W}(T)$. This will follow if it is shown that any closed half-plane which contains $\sigma(T)$ also contains $\overline{W}(T)$. By translation and rotation this reduces to showing that

$$\operatorname{Re} \sigma(T) \leq 0 \text{ implies } \operatorname{Re} \overline{W}(T) \leq 0.$$

Let $\|x\| = 1$ and $Tx = (a+ib)x + y$ with a, b real and x orthogonal to y. Then from V

$$\|(T-cI)x\| \geq \operatorname{dist}(c, \sigma(T)) \geq c$$

for all $c > 0$, so that

$$c^2 \leq \|(T-cI)x\|^2 = (a-c)^2 + b^2 + \|y\|^2$$

and therefore $2ac \leq a^2 + b^2 + \|y\|^2$. Since this holds for all $c > 0$, $\operatorname{Re}(Tx, x) = a \leq 0$ follows.

REMARKS. 1. Halmos [20] and Bram [9] have shown that the following condition characterizes subnormality: for every integer $n \geq 1$ and choice of vectors $x_1, x_2, \ldots, x_n$, the matrix

$$(T^i x_j, T^j x_i)$$

is positive definite.

2. For further information on hyponormality the reader may consult the papers of Stampfli, from which many of the above results have been taken.

3. The important classes of Laurent and Toeplitz operators are defined as follows. Let $L^2 = L^2(\lambda)$, where λ is Lebesgue measure on the unit circle, and let H^2 be the subspace of $f \in L^2$ all of whose negative Fourier coefficients vanish. The Laurent operators are those of the form $M_\phi f = \phi f$ for $f \in L^2$, where $\phi \in L^\infty$, and the Toeplitz operators are the compressions $T_\phi = PM_\phi | H^2$, where P is the projection on H^2. Basic information on these operators is contained in [13] and [24, Ch. 20].

4. The class of operators A such that A and A^*A commute is introduced and studied by Brown in [12].

Exercises. 1. If A is an operator and $\mathfrak{M}$ an invariant subspace, then $(A|\mathfrak{M})^* = PA^*|\mathfrak{M}$, where P is the projection (orthogonal) on $\mathfrak{M}$. If A is normal, then $A|\mathfrak{M}$ is normal if and only if $\mathfrak{M}$ reduces.

2. If T is hyponormal and $\mathfrak{M}$ an invariant subspace, then $T|\mathfrak{M}$ is hyponormal. If $T|\mathfrak{M}$ is normal, then $\mathfrak{M}$ reduces.

3. The span of the eigenvectors of a hyponormal operator T is a reducing subspace in which T is normal.

4. If T is hyponormal and quasinilpotent then $T = 0$. (T is called quasinilpotent if $r(T) = 0$.)

5. An isolated point of the spectrum of a hyponormal operator is an eigenvalue.

6. A compact hyponormal operator is normal.

7. If T is hyponormal and $\sigma(T) \subset \{z \mid |z| = 1\}$, then T is unitary. (Hint: both T and T^{-1} are contractions.) If $\sigma(T)$ is real then T is self-adjoint.

8. Every invariant subspace of a compact normal operator reduces. If every invariant subspace of a compact operator C reduces, then C is normal. (Hint: there is a minimal invariant subspace $\mathfrak{M}$ such that $\|C|\mathfrak{M}\| = \|C\|$.)

SECTION 1.

SHIFTS

The shift operators are of fundamental importance in many parts of operator theory. In this section they are introduced with some of their basic properties.

Let $\mathcal{K}$ be a Hilbert space, and let $\ell^2_+(\mathcal{K}) = \mathcal{K} \oplus \mathcal{K} \oplus \cdots$ be the Hilbert space of all sequences $x = \{x_n\}_{n=0}^\infty$ of vectors $x_n \in \mathcal{K}$ such that $\|x\|^2 = \Sigma_{n=0}^\infty \|x_n\|^2 < \infty$. The *unilateral shift operator* U_+ on $\ell^2_+(\mathcal{K})$ is defined by

$$U_+(x_0, x_1, \ldots) = (0, x_0, x_1, \ldots) \ .$$

The *multiplicity* of U_+ is the cardinal number $n = \dim \mathcal{K}$. It is easy to see that $U_+^*(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$ (this operator is called the *backward shift*), and that unilateral shift operators are unitarily equivalent if and only if they have the same multiplicity.

Recall that an operator V is an *isometry* if $\|Vx\| = \|x\|$ for all vectors x. This is equivalent to $(Vx, Vy) = (x, y)$ for all x and y, and to $V^*V = I$.

I. The operator V on $\mathcal{H}$ is unitarily equivalent to a unilateral shift operator if and only if V is an isometry with

$$\bigcap_{n=0}^{\infty} V^n \mathcal{H} = \{0\} \ .$$

The multiplicity is $\dim (V\mathcal{H})^\perp$.

Proof: Let V be such an isometry, and let $\mathcal{K}_0 = (V\mathcal{H})^{\perp}$ and $\mathcal{K}_n = V^n \mathcal{K}_0$ for $n \geq 1$. It will be shown that these spaces are mutually orthogonal and span $\mathcal{H}$. Since $\mathcal{K}_n \subset V\mathcal{H}$ for $n \geq 1$, we have $\mathcal{K}_0 \perp \mathcal{K}_n$ for $n \geq 1$. Hence

$$\mathcal{K}_\ell = V^\ell \mathcal{K}_0 \perp V^\ell \mathcal{K}_n = \mathcal{K}_{\ell+n}$$

for $n \geq 1$ and $\ell \geq 0$. Thus $\{\mathcal{K}_n\}^{\perp}$. Suppose $z \perp \mathcal{K}_n$ for all $n \geq 0$. Now $\mathcal{H} = V\mathcal{H} \oplus \mathcal{K}_0$, so $V\mathcal{H} = V^2\mathcal{H} \oplus V\mathcal{K}_0 = V^2\mathcal{H} \oplus \mathcal{K}_1$, and therefore $\mathcal{H} = V^2\mathcal{H} \oplus \mathcal{K}_1 \oplus \mathcal{K}_0$. Repeating gives

$$\mathcal{H} = V^n\mathcal{H} \oplus \mathcal{K}_{n-1} \oplus \cdots \oplus \mathcal{K}_0 \text{ for all } n \geq 1 .$$

Consequently $z \in V^n\mathcal{H}$ for $n \geq 0$, and so $z = 0$.

Now let U_+ be the shift on $\ell^2_+(\mathcal{K}_0)$, and define $W : \ell^2_+(\mathcal{K}_0) \to \mathcal{H}$ by $W(x_0, x_1, x_2, \dots) = x_0 + Vx_1 + V^2x_2 + \cdots$. Then W is evidently unitary from $\ell^2_+(\mathcal{K}_0)$ onto $\mathcal{H}$, and

$$VW(x_0, x_1, x_2, \dots) = Vx_0 + V^2x_1 + \cdots$$

$$WU_+(x_0, x_1, x_2, \dots) = W(0, x_0, x_1, \dots) = Vx_0 + V^2x_1 + \cdots,$$

so that $W^*VW = U_+$.

Such operators are called *pure isometries,* or simply unilateral shift operators. The next result is referred to by many authors as the Wold decomposition.

II. Structure of Isometries. [37, §X; 22] If V is an isometry, there is a unique reducing subspace $\mathfrak{M}$ such that $V|\mathfrak{M}$ is unitary and $V|\mathfrak{M}^{\perp}$ is a pure isometry.

Proof: Let $\mathfrak{M} = \cap_{n=0}^{\infty} V^n\mathcal{H}$. Then $\mathfrak{M}$ is an invariant subspace and if $x \perp \mathfrak{M}$, $Vx \perp V\mathfrak{M} = \cap_{n=1}^{\infty} V^n\mathcal{H} = \mathfrak{M}$, so $\mathfrak{M}$ reduces V. Next, $V|\mathfrak{M}$ is an isometry and $V\mathfrak{M} = \mathfrak{M}$, so $V|\mathfrak{M}$ is unitary. Finally, $V|\mathfrak{M}^{\perp}$ is an isometry and $\cap_{n=0}^{\infty} V^n(\mathfrak{M}^{\perp}) = \{0\}$, so $V|\mathfrak{M}^{\perp}$ is a shift by I.

Let $\mathcal{K}$ be a Hilbert space, let $\ell^2(\mathcal{K}) = \Sigma_{-\infty}^{\infty} \oplus \mathcal{K}$ be the Hilbert space of two-way sequences $x = (\ldots, x_{-1}, \hat{x}_0, x_1, \ldots)$ of vectors from $\mathcal{K}$ with $\|x\|^2 = \Sigma_{-\infty}^{\infty} \|x_n\|^2 < \infty$, and define the *bilateral shift U on* $\ell^2(\mathcal{K})$ by

$$U(\ldots, x_{-1}, \hat{x}_0, x_1, \ldots) = (\ldots, \hat{x}_{-1}, x_0, x_1, \ldots) .$$

The multiplicity of U is dim $\mathcal{K}$. Note that U is unitary.

III. If V is an isometry on $\mathcal{H}$, there is a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a unitary operator U on $\mathcal{K}$ such that $U\mathcal{H} \subset \mathcal{H}$ and $U|\mathcal{H} = V$.

Proof: Use II and extend the unilateral shift summand to the corresponding bilateral shift.

DEFINITION. A subspace $\mathfrak{M}$ of $\mathcal{H}$ is *wandering* for the unitary operator W on $\mathcal{H}$ if the subspaces $\{W^n\mathfrak{M}\}_{-\infty}^{\infty}$ are pairwise orthogonal, and *complete* if they span $\mathcal{H}$.

LEMMA. If $\mathfrak{M}$ is a wandering subspace for the bilateral shift U on $\ell^2(\mathcal{K})$, then $\dim \mathfrak{M} \leq \dim \mathcal{K}$.

Proof: (Halperin). If $\mathcal{K}$ is infinite-dimensional, then $\dim \mathfrak{M} \leq \dim \ell^2(\mathcal{K}) = \dim \mathcal{K}$, as required. So let $x_1, \ldots, x_k$ be an orthonormal basis of $\mathcal{K}$; then $\{U^n x_i \mid n = 0, \pm 1, \ldots; 1 \leq i \leq k\}$ is an orthonormal basis of $\ell^2(\mathcal{K})$. If $\{y_a\}$ is an orthonormal basis for $\mathfrak{M}$, then $\{U^n y_a\}$ is an orthonormal set in $\ell^2(\mathcal{K})$, so $\|x_i\|^2 \geq \Sigma_{n,a} |(x_i, U^n y_a)|^2$ by Bessel's inequality. Hence

$$\dim \mathcal{K} = \sum_i \|x_i\|^2 \geq \sum_{n,a,i} |(x_i, U^n y_a)|^2$$

$$= \sum_{n,a,i} |(U^n x_i, y_a)|^2 = \sum_a \|y_a\|^2 = \dim \mathfrak{M} .$$

IV. Bilateral shifts are unitarily equivalent if and only if they have the same multiplicity.

V. The operator W on $\mathcal{H}$ is unitarily equivalent to a bilateral shift if and only if it is unitary and has a complete wandering subspace $\mathfrak{M}$. The multiplicity is $\dim \mathfrak{M}$.

We conclude this section with a result like II.

LEMMA. Let W be unitary on $\mathcal{H}$, let $\mathfrak{M}_0$ be an invariant subspace, and let $\mathfrak{M}_k = W^{-k}(\mathfrak{M}_0)$ for $k = 0, \pm 1, \dots$. Then

1) $\mathfrak{M}_k \subset \mathfrak{M}_n$ for all $k \leq n$,

2) the spaces $\mathfrak{M}_{-\infty} = \cap \mathfrak{M}_k$ and $\mathfrak{M}_\infty = \overline{\cup \mathfrak{M}_k}$ reduce W, and

3) $W|(\mathfrak{M}_\infty \ominus \mathfrak{M}_{-\infty})$ is a bilateral shift of multiplicity $\dim(\mathfrak{M}_1 \ominus \mathfrak{M}_0)$.

Proof: If $k \leq n$, then $W^{n-k}\mathfrak{M}_0 \subset \mathfrak{M}_0$ and consequently $\mathfrak{M}_k = W^{-k}\mathfrak{M}_0 \subset W^{-n}\mathfrak{M}_0 = \mathfrak{M}_n$. Since W shifts the family $\{\mathfrak{M}_k\}$ one step back, and $W^* = W^{-1}$ shifts it one step forward, it is clear that $\mathfrak{M}_\infty$ and $\mathfrak{M}_{-\infty}$ are reducing. Finally,

$$\mathfrak{M}_\infty \ominus \mathfrak{M}_{-\infty} = \sum_k \oplus (\mathfrak{M}_{k+1} \ominus \mathfrak{M}_k),$$

and

$$W^{-k}(\mathfrak{M}_1 \ominus \mathfrak{M}_0) = \mathfrak{M}_{k+1} \ominus \mathfrak{M}_k,$$

so $W|(\mathfrak{M}_\infty \ominus \mathfrak{M}_{-\infty})$ is a bilateral shift by V.

Notice that $\mathfrak{M}_0$ reduces W if and only if $\mathfrak{M}_\infty = \mathfrak{M}_{-\infty}$.

VI. If W is unitary on $\mathcal{H}$, there is a reducing subspace $\mathfrak{M}$ such that $W|\mathfrak{M}$ is a bilateral shift and such that every invariand subspace of $W|\mathfrak{M}^\perp$ is reducing.

Proof: Consider a maximal orthogonal family $\{\mathfrak{M}_\alpha\}$ of reducing subspaces such that each $W|\mathfrak{M}_\alpha$ is a bilateral shift. If $\mathfrak{M} = \sum \oplus \mathfrak{M}_\alpha$, then by the lemma every invariant subspace of $W|\mathfrak{M}^\perp$ reduces.

Exercises. 1. If U is a shift (unilateral or bilateral) of multiplicity n, then U^k is a shift of multiplicity kn. If U_α is a shift of multiplicity n_α, then $\sum \oplus U_\alpha$ is a shift of multiplicity $\sum n_\alpha$.

2. Define T on $L^2[0,1]$ by $(Tf)(t) = \sqrt{2}f(2t)$ on $[0,\frac{1}{2}]$ and $(Tf)(t) = 0$ on $[\frac{1}{2},1]$. Then T is a unilateral shift of infinite multiplicity.

3. The operator $(Tf)(t) = \sqrt{2}f(2t)$ on $L^2[0,\infty]$ is a bilateral shift.

4. Carry out the decomposition VI for the operator W on $L^2(\mu)$ defined by $(Wf)(\theta) = e^{i\theta}f(\theta)$, where μ is a positive Borel measure on the unit circle.

5. Is the space $\mathfrak{M}$ of VI unique?

REMARKS. [24] contains basic information on shifts. Many of the results here and in Sections 3 and 4 can be found in [22]. For material related to VI see [19] and [57].

SECTION 2.

MODELS

If $\mathfrak{M}$ is an invariant subspace for the unilateral shift U_+, then $U_+|\mathfrak{M}$ is again a unilateral shift (by 1. I). Since $\mathfrak{M}^\perp$ is invariant for U_+^*, it is natural to ask what can be said about the operators $U_+^*|\mathfrak{M}^\perp$. The remarkable answer, due to Rota [41], is provided by the following construction.

Let T be an operator on a Hilbert space $\mathcal{K}$ with $\|T\| < 1$. Define a map $S: \mathcal{K} \to \ell_+^2(\mathcal{K})$ by $Sx = (x, Tx, T^2x, \ldots)$. Then S is linear, one-to-one, and bounded, since

$$\|Sx\|^2 = \Sigma\|T^n x\|^2 \leq \Sigma\|T\|^{2n}\|x\|^2 = (1-\|T\|^2)^{-1}\|x\|^2 .$$

It is clearly bounded below, and so $\mathfrak{R}$ = range S is a closed subspace of $\ell_+^2(\mathcal{K})$. Moreover $ST = U_+^*S$, and therefore $\mathfrak{R}$ is invariant for U_+^*, and the operators T and $U_+^*|\mathfrak{R}$ are similar.

Thus any strict contraction is similar to an operator of the form $U_+^*|\mathfrak{M}$. With a little more care the argument can be made to yield considerably more. The hypothesis $\|T\| < 1$ was used only to ensure that S is bounded, but for this it is obviously sufficient to have the convergence of the series $\Sigma\|T^n\|^2$. The latter will be the case if $\overline{\lim}\,\|T^n\|^{2/n} < 1$. Since $\lim\|T^n\|^{1/n}$ always exists and equals the spectral radius $r(T)$, the following has been proved;

I. Rota's Theorem. Any operator with spectral radius less than one is similar to a part of the backward shift.

COROLLARY 1. For any operator T, $r(T) = \inf \|S^{-1}TS\|$, the infimum being taken over all invertible operators S.

Proof: For any $\varepsilon > 0$ the operator $(r(T) + \varepsilon)^{-1}T$ has spectral radius less than one, and an application of the theorem leads to the conclusion that $r(T) \geq \inf \|S^{-1}TS\|$. The other inequality follows from $r(T) = r(S^{-1}TS) \leq \|S^{-1}TS\|$.

COROLLARY 2. (S. Hildebrandt). For any operator T,

$$\mathrm{co}(\sigma(T)) = \cap\{\overline{W}(S^{-1}TS) \mid S \text{ invertible}\},$$

where co denotes convex hull and $\overline{W}$ the closure of the numerical range.

Proof: For any operator A, $\overline{W}(A)$ is convex and contains $\sigma(A)$ [24]. Then $\sigma(T) = \sigma(S^{-1}TS) \subset \overline{W}(S^{-1}TS)$, and one inclusion follows. The proof of the other inclusion is based on the simple fact that any closed convex subset of the plane is the intersection of all open discs containing it. Let D be an open disc containing $\mathrm{co}(\sigma(T))$ with center λ and radius r. Then the spectrum of $(1/r)(T - \lambda I)$ lies in the open unit disc, so by Corollary 1 there is an invertible operator S such that $\|(1/r)S^{-1}(T - \lambda I)S\| < 1$. It follows that $\overline{W}((1/r)S^{-1}(T - \lambda I)S)$ is contained in the open unit disc, or equivalently that $\overline{W}(S^{-1}TS) \subset D$. This proof is due to J. P. Williams.

From the point of view of model theory, the above construction suffers the defect that, among all possible ways of representing a given operator as a part of a backward shift, it it may not give the simplest. For example, the parts of the shift of multiplicity one which arise in this way are all one-dimensional. The following refinement of Rota's technique, due to de Branges and Rovnyak [11, Appendix], helps to clear up this objection, and at the same time replaces similarity by unitary equivalence.

II. Theorem. The operator T on $\mathcal{H}$ is unitarily equivalent to a part of the backward shift if and only if $\|T\| \leq 1$ and $T^n \to 0$ strongly.

Proof: Let T be a contraction such that $T^n \to 0$ strongly. Let R be the non-negative square root of $I - T^*T$ (notice that $I - T^*T \geq 0$ since $\|T\| \leq 1$), let $\mathcal{K} = (\text{range } R)^-$, and define $W: \mathcal{H} \to \ell^2_+(\mathcal{K})$ by

$$Wx = (Rx, RTx, RT^2x, \ldots) \ .$$

Then $\|RT^kx\|^2 = (T^{*k}R^2T^kx, x) = \|T^kx\|^2 - \|T^{k+1}x\|^2$, so

$$\sum_{k=0}^{n} \|RT^nx\|^2 = \|x\|^2 - \|T^{n+1}x\|^2 \to \|x\|^2 \ .$$

Therefore W is unitary from $\mathcal{H}$ onto $\mathfrak{R} = \text{range } W$. Obviously, $WT = U_+^*W$, and consequently $\mathfrak{R}$ is invariant for U_+^*, and T and $U_+^* \mid \mathfrak{R}$ are unitarily equivalent.

COROLLARY. If $\|T\| < 1$, then T is unitarily equivalent to a part of the backward shift.

This procedure gets by with the shift of smallest possible multiplicity. For example, applying it with $T = U_+^*$, one gets W = identity. Again, T is a part of the backward shift of multiplicity n if and only if $I - T^*T$ is of rank at most n.

These results lead to the following reformulation of the invariant subspace problem: are the minimal non-zero invariant subspaces of the backward shift one-dimensional? This has been verified for shifts of finite multiplicity.

Exercise. (Sarason [43]). If V is the Volterra operator defined on $L^2[0,1]$ by $(Vf)(t) = \int_0^t f$, then $T = (I - V)(I + V)^{-1}$ is unitarily equivalent to a part of the backward shift of multiplicity one.

SECTION 3

INVARIANT SUBSPACES OF SIMPLE SHIFTS

The results of the previous section make the nature of the invariant subspaces of shift operators a matter of great interest. The reducing subspaces will be investigated first, after establishing a useful alternative representation of the shifts. Much of the material in this and the next section can be found in Helson [28].

I. The functions $e_n(\theta) = e^{in\theta}$, $n \in \mathbf{Z}$, form a complete orthonormal set in $L^2(\lambda) = L^2$, where λ is normalized Lebesgue measure on the unit circle.

Proof: By elementary calculus the set $\{e_n\}$ is orthonormal. For any continuous function f there is a sequence $\{p_n\}$ of trigonometric polynomials (linear combinations of the functions e_n) such that $p_n \to f$ uniformly (Stone-Weierstrass Theorem). Then $p_n \to f$ in the L^2 norm since

$$\|p_n - f\|_2 \leq \max |p_n - f| \quad .$$

Thus the closed linear span of the e_n contains the continuous functions. Since the latter are $\| \ \|_2$-dense in L^2 the result follows.

Let H^2 be the closed linear span of the e_n with $n \geq 0$. Notice that H^2 consists of the $f \in L^2$ all of whose negative Fourier coefficients vanish: $\int_0^{2\pi} f(\theta)e^{in\theta}d\theta = 0$ for all $n > 0$.

For $\phi \in L^\infty$, consider the operator M_ϕ defined on L^2 by $M_\phi f = \phi f$. It is easily seen that $\|M_\phi\| = \|\phi\|_\infty$. The next two results, which are concerned with the operator M_e (where $e(\theta) = e_1(\theta) = e^{i\theta}$), follow from V and I of §1.

II. M_e is unitarily equivalent to the bilateral shift of multiplicity one.

III. $M_e | H^2$ is unitarily equivalent to the unilateral shift of multiplicity one.

IV. An operator on L^2 commutes with the bilateral shift M_e if and only if it is of the form M_ϕ for some $\phi \in L^\infty$.

Proof: Suppose that A commutes with M_e, and let $\phi = Ae_0$. It will be shown that $\phi \in L^\infty$ and that $A = M_\phi$. Since M_e is unitary, A also commutes with $M_e^* = M_{\bar{e}}$. It follows that A commutes with M_p for any trigonometric polynomial p, and hence that

$$Ap = AM_p e_0 = M_p A e_0 = M_p \phi = M_\phi p \quad .$$

The rest of the argument amounts to showing that the linear transformation M_ϕ is closed (it has not yet been shown that $\phi \in L^\infty$). If $f \in L^2$, by I there is a sequence $\{p_n\}$ of trigonometric polynomials such that $p_n \to f$ in the L^2 norm, so that

$$\phi p_n = M_\phi p_n = Ap_n \to g \ ,$$

where $g = Af$. Now let $\delta > 0$ and let $S = \{\theta \mid |\phi(\theta)| \geq \delta\}$. Then

$$\int_0^{2\pi} |\phi p_n - g|^2 \geq \int_S |\phi p_n - g|^2$$

$$= \int_S |\phi|^2 \, |p_n - (g/\phi)|^2 \geq \delta^2 \int_S |p_n - (g/\phi)|^2$$

so that $p_n \to g/\phi$ in $L^2(S)$. But $p_n \to f$ in L^2 and therefore in $L^2(S)$, and so $\phi f = g$ a.e. in S. By the arbitrary nature of δ this gives $\phi f = g$ a.e. in $\{\theta \,|\, \phi(\theta) \neq 0\}$. From $\phi p_n \to g$ follows $g = 0$ a.e. in $\{\theta \,|\, \phi(\theta) = 0\}$, and hence $Af = g = \phi f = M_\phi f$. Thus $A = M_\phi$.

Letting χ be the characteristic function of S, we now have

$$\|A\chi\|^2 = \int_0^{2\pi} |\phi\chi|^2 \geq \delta^2 \lambda(S) = \delta^2 \|\chi\|^2 ,$$

so $\chi = 0$ for $\delta > \|A\|$ and $\|\phi\|_\infty \leq \|A\|$. This completes the proof. For another proof see §4.IV.

COROLLARY. A subspace of L^2 reduces the bilateral shift if and only if it is of the form

$$\{ f \,|\, f \in L^2 \text{ and } f = 0 \text{ a.e. on } S\}$$

for some measurable subset S of the circle.

Proof: The projection on a reducing subspace commutes with M_e and consequently is of the form M_ϕ for some $\phi \in L^\infty$. Since M_ϕ is a projection, $\phi^2 = \phi$ and $\phi(\theta) = 0$ or 1 a.e. Let $S = \{\theta \,|\, \phi(\theta) = 0\}$.

V. The unilateral shift of multiplicity one is irreducible.

Proof: Let $\mathfrak{M}$ be a subspace of H^2 which reduces $U_+ = M_e | H^2$. Then $U_+ H^2 = U_+ \mathfrak{M} \oplus U_+ \mathfrak{M}^\perp \subset U_+ \mathfrak{M} \oplus \mathfrak{M}^\perp$, so that $\mathfrak{M} \ominus U_+ \mathfrak{M} \subset H^2 \ominus U_+ H^2$. With $\mathfrak{L} = \mathfrak{M} \ominus U_+ \mathfrak{M}$, it follows that either $\mathfrak{L} = \{0\}$ or $\mathfrak{L} = H^2 \ominus U_+ H^2$. But $U_+ | \mathfrak{M}$ is a unilateral shift by §1.I, so that

$$\mathfrak{M} = \mathfrak{L} \oplus U_+\mathfrak{L} \oplus U_+^2\mathfrak{L} \oplus \cdots$$

as in the proof of that result, and consequently $\mathfrak{M} = \{0\}$ or $\mathfrak{M} = H^2$.

The invariant subspace situation, which is much more complicated, was investigated by Beurling [8].

VI. The non-zero invariant non-reducing subspaces of the bilateral shift in L^2 are those of the form qH^2, with q a measurable function such that $|q| = 1$ a.e. The function q is demined by the subspace up to a constant factor.

Proof: If $\mathfrak{M}$ is invariant and non-reducing, then $e\mathfrak{M} \subsetneqq \mathfrak{M}$. (This is true of any isometry.) Let q be a unit vector with $q \in \mathfrak{M}$, $q \perp e\mathfrak{M}$. Then $q \perp e_n q$ for $n \geq 1$: $\int_0^{2\pi} |q(\theta)|^2 e^{in\theta} d\theta = 0$ for $n \geq 1$. Conjugating, the same result holds for all $n \neq +$, so $|q|$ is constant a.e. Since $\|q\| = 1$, this constant is 1.

The set $\{qe_n\}_{n \in \mathbf{Z}}$ is therefore orthonormal, and its span reduces the shift. By the Corollary to IV the span is all of L^2. Since $qe_n \in \mathfrak{M}$, $n \geq 0$, we have $qH^2 \subset \mathfrak{M}$. On the other hand, $(qH^2)^\perp$ is spanned by the qe_n with $n < 0$. But $q \perp e\mathfrak{M}$ implies $q \perp e_n\mathfrak{M}$ for all $n > 0$, or equivalently, $qe_n \perp \mathfrak{M}$ for all $n < 0$, and so $(qH^2)^\perp \subset \mathfrak{M}^\perp$. Therefore $qH^2 = \mathfrak{M}$.

For uniqueness, suppose $pH^2 = qH^2$. Then $\bar{p}qH^2 = H^2$, so $\bar{p}q \in H^2$. Similarly $p\bar{q} \in H^2$, and therefore $\bar{p}q = q/p$ is constant.

COROLLARY. If $f \in H^2$ vanishes on a set of positive measure, then $f = 0$ a.e.

Proof: Suppose it is not true that $f = 0$ a.e. Then the closed linear span $\mathfrak{M}_f$ of the functions fe_n $(n \geq 0)$ is non-zero, invariant, and non-reducing (since $e_{-n} f \notin H^2$ for sufficiently large n). Therefore $\mathfrak{M}_f = qH^2$ for some q of modulus one, by VI. This is impossible, since there is a set of positive measure on which all members of $\mathfrak{M}_f$ vanish.

VII. The non-zero invariant subspaces of the unilateral shift in H^2 are those of the form qH^2, with $q \in H^2$ and $|q| = 1$ a.e. The function q is determined by the subspace up to a constant factor.

Proof: If $\mathfrak{M}$ is invariant for the unilateral shift, then viewed as a subspace of L^2 it is invariant for the bilateral shift, so $\mathfrak{M} = qH^2$ for some $q \in L^2$ with $|q| = 1$ a.e. But $\mathfrak{M} \subset H^2$, so $q \in H^2$.

The structure of functions $q \in H^2$ with $|q| = 1$ a.e. is well-understood [30, Ch. 5]. Let $q = \Sigma_{n=0}^{\infty} a_n e_n$. The corresponding functions $\hat{q}(z) = \Sigma_{n=0}^{\infty} a_n z^n$, analytic in the open unit disc, are called *inner functions*. Each inner function is a product $c \cdot B \cdot S$, where

1) c is a constant of modulus 1;

2) B is a Blaschke product:

$$B(z) = z^k \prod_{n=1}^{\infty} \frac{a_n - z}{1 - \bar{a}_n z} \cdot \frac{\bar{a}_n}{|a_n|}$$

where k is a non-negative integer, and the a_n are complex numbers satisfying $0 < |a_n| < 1$ and $\Sigma(1 - |a_n|) < \infty$;

3) S is a singular function:

$$S(z) = \exp\left\{-\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta)\right\}$$

where μ is a positive finite Borel measure on the circle which is singular with respect to Lebesgue measure.

Exercise. Let $q \in H^2$ be of modulus one a.e., let $\hat{q}$ be defined as above, and let

$$k(e^{i\theta}, z) = \mathrm{Re}\left(\frac{e^{i\theta}+z}{e^{i\theta}-z} \right)$$

for $|z| < 1$. Show that

$$\hat{q}(z) = \frac{1}{2\pi} \int_0^{2\pi} q(e^{i\theta})\, k(e^{i\theta}, z)\, d\theta \quad .$$

Use the properties $k \geq 0$ and

$$\frac{1}{2\pi} \int_0^{2\pi} k(e^{i\theta}, z)\, d\theta = 1$$

of the kernel k, valid for all θ and $|z| < 1$, to show that $|\hat{q}(z)| \leq 1$ for $|z| < 1$.

SECTION 4.

INVARIANT SUBSPACES OF GENERAL SHIFTS

Shifts of arbitrary countable multiplicity will be considered in this section. The procedure is much the same as in the preceding section, except that it will be necessary to use direct sums of the L^2-spaces of the circle. Specifically, let $\mathcal{K}$ be a separable Hilbert space, and consider functions defined on the unit circle with values in $\mathcal{K}$. Such a function f will be called *measurable* if the complex-valued function $(f(\cdot), x)$ on the circle is measurable in the usual sense for each $x \in \mathcal{K}$.

The separability of $\mathcal{K}$ is used in the following lemma.

I. Lemma. If f and g are measurable, so is $(f(\cdot), g(\cdot))$.

Proof: If $\{b_\alpha\}$ is an orthonormal basis of $\mathcal{K}$, then $f(\theta) = \Sigma_\alpha (f(\theta), b_\alpha) b_\alpha$, and so $(f(\theta), g(\theta)) = \Sigma_\alpha (f(\theta), b_\alpha)(b_\alpha, g(\theta))$ is measurable.

DEFINITION. $L^2(\mathcal{K})$ consists of all measurable functions f on the unit circle to $\mathcal{K}$ such that

$$\|f\|^2_{L^2(\mathcal{K})} = \frac{1}{2\pi} \int_0^{2\pi} \|f(\theta)\|^2_{\mathcal{K}} \, d\theta < \infty ,$$

functions which differ only on a set of measure zero being identified. This is a Hilbert space with respect to the inner product

$$(f,g)_{L^2(\mathcal{K})} = \frac{1}{2\pi}\int_0^{2\pi} (f(\theta),g(\theta))_{\mathcal{K}}\,d\theta$$

(completeness will be clear presently).

Functions in $L^2(\mathcal{K})$ admit two kinds of orthogonal expansion, as follows.

1) Let $\{b_a\}$ be an orthonormal basis of $\mathcal{K}$. If $f \in L^2(\mathcal{K})$, the functions $f_a(\theta) = (f(\theta), b_a)$ are called the *coordinate functions* of f. Of course, $f(\theta) = \Sigma\, f_a(\theta)b_a$ (convergence in the norm of $\mathcal{K}$. The second of these equations shows that the coordinate functions are in L^2. In fact,

$$\|f\|^2_{L^2(\mathcal{K})} = \frac{1}{2\pi}\int_0^{2\pi} \|f(\theta)\|^2_{\mathcal{K}}\,d\theta = \sum_a \frac{1}{2\pi}\int_0^{2\pi} |f_a(\theta)|^2 d\theta$$

$$= \sum_a \|f_a\|^2_{L^2} \quad .$$

Thus the map $f \to \Sigma_a \oplus f_a$ is an isometry from $L^2(\mathcal{K})$ into $\Sigma_a \oplus L^2$. To see that it is actually onto, let $\Sigma \oplus f_a \in \Sigma \oplus L^2$. Then

$$\Sigma \|f_a\|^2_{L^2} = \frac{1}{2\pi}\int_0^{2\pi} \Sigma\, |f_a(\theta)|^2\, d\theta$$

if finite, so that $\Sigma\, |f_a(\theta)|^2 < \infty$ a.e., and therefore the function $f(\theta) = \Sigma\, f_a(\theta)b_a$ is defined a.e. It is measurable, and since

$$\frac{1}{2\pi}\int_0^{2\pi} \|f(\theta)\|^2_{\mathcal{K}}\,d\theta = \Sigma \|f_a\|^2_{L^2} < \infty ,$$

it is in $L^2(\mathcal{K})$. This proves

II. Lemma. The map $f \to \Sigma \oplus f_a$ is an isometry from $L^2(\mathcal{K})$ onto $\Sigma_a \oplus L^2$. Therefore $L^2(\mathcal{K})$ is complete.

2) The second kind of expansion is as a Fourier series with vector coefficients.

III. Lemma. Each $f \in L^2(\mathcal{K})$ admits a unique expansion

$$f(\theta) = \sum_{-\infty}^{+\infty} x_k e^{ik\theta},$$

with $x_k \in \mathcal{K}$ and $\|f\|^2_{L^2(\mathcal{K})} = \Sigma \|x_k\|^2_{\mathcal{K}}$. This is to be understood in the weak sense: for each $x \in \mathcal{K}$, the Fourier expansion of $(f(\,\cdot\,), x)$ is $\Sigma (x_k, x) e^{ik\theta}$.

Proof: $x \to 1/2\pi \int_0^{2\pi} (f(\theta), x) e^{-ik\theta} d\theta$ is a bounded conjugate linear functional on $\mathcal{K}$, so there exists a unique $x_k \in \mathcal{K}$ such that

$$(x_k, x) = \int_0^{2\pi} (f(\theta), x) e^{ik\theta}\, d\theta \quad .$$

Hence $\Sigma (x_k, x) e^{ik\theta}$ is the Fourier expansion of $(f(\theta), x)$, $\|(f(\,\cdot\,), x)\|^2 = \Sigma_k |(x_k, x)|^2$, and

$$\|f\|^2 = \sum_a \|f_a\|^2 = \sum_a \sum_k |(x_k, b_a)|^2 = \sum \|x_k\|^2 \,.$$

COROLLARY. The map $f \to \{x_k\}$ is an isometry from $L^2(\mathcal{K})$ onto $\ell^2(\mathcal{K})$.

Multiplication by a bounded measurable complex-valued function defines a bounded operator in $L^2(\mathcal{K})$. In particular, it follows easily from the foregoing Fourier expansion that multiplication by e is the bilateral shift U of multiplicity dim $\mathcal{K}$. The subspace

$$H^2(\mathcal{K}) = \{f \mid x_k = 0 \text{ for all } k < 0\}$$

is invariant under U, and $U_+ = U \mid H^2(\mathcal{K})$ is the unilateral shift of multiplicity dim $\mathcal{K}$. Notice that $f \in H^2(\mathcal{K})$ if and only if $f_\alpha \in H^2$ for all α. Another helpful observation is that $\mathcal{K}$ can be isometrically embedded in $H^2(\mathcal{K})$, by identifying $x \in \mathcal{K}$ with the function identically equal to x. This identification will often be made in order to simplify the notation.

Invariant subspaces will now be investigated in terms of this realization of the shift operators. As before, the reducing subspaces will be treated first, by means of a characterization of all operators which commute with the shift. Of course multiplication by any L^∞ function commutes with the shift, but if dim $\mathcal{K} > 1$ there are others. For example, if dim $\mathcal{K} = 2$, then any operator defined on $L^2(\mathcal{K}) = L^2 \oplus L^2$ by means of a 2×2 matrix of L^∞ functions commutes with the shift. Such a matrix may be viewed as a function on the circle with values in the bounded operators on $\mathcal{K}$. The burden of the following theorem is that every commuting operator is of this type.

DEFINITION. An operator-valued function A: circle $\to$ $\mathcal{B}(\mathcal{K})$ is *measurable* if $A(\cdot)x$ is measurable for all $x \in \mathcal{K}$. If in addition the function $\|A(\cdot)\|$ is essentially bounded, one can define a bounded operator $\hat{A}$ on $L^2(\mathcal{K})$ as follows:

$$(\hat{A}f)(\theta) = A(\theta)f(\theta) \quad .$$

It is easily seen that $\hat{A}f$ is measurable, and

$$\|\hat{A}f\|^2 = \frac{1}{2\pi}\int_0^{2\pi} \|(\hat{A}f)(\theta)\|^2 d\theta$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \|A(\theta) f(\theta)\|^2 d\theta$$

$$\leq (\text{ess sup } \|A(\theta)\|^2) \cdot \|f\|^2 \quad ,$$

so $\|\hat{A}\| \leq \text{ess sup } \|A(\theta)\|$. Observe that $\hat{A}$ commutes with multiplications, and in particular with U.

IV. Theorem. If $\mathbf{A}$ is an operator in $L^2(\mathcal{K})$ that commutes with the bilateral shift U, there exists an essentially unique essentially bounded measurable operator-valued function A such that $\mathbf{A} = \hat{A}$.

The proof will be reduced to the case of a unitary operator $\mathbf{A}$, by means of:

LEMMA. If the operator A commutes with the normal operator N, then A is a linear combination of four unitary operators each of which commutes with N.

Proof: Let $A = R + iS$ with R and S self-adjoint. Then R and S commute with N, since N is normal. It can be assumed that R and S are contractions, and then $R \pm i(I - R^2)^{1/2}$ and $iS \pm (I - S^2)^{1/2}$ are unitary, commute with N, and have sum $2A$.

Proof of Theorem: By the lemma it will suffice to consider a unitary operator $\mathbf{A}$ that commutes with the shift. Let $\{b_a\}$ be an orthonormal basis of $\mathcal{K}$, and g_a a function in the equivalence class $\mathbf{A}b_a$ for each a. Define $A(\theta)b_a = g_a(\theta)$. If $\phi \in L^\infty$

$$(\mathbf{A}(\phi b_a), \mathbf{A} b_\beta) = (\phi b_a, b_\beta) = \frac{1}{2\pi} \int_0^{2\pi} \phi(\theta)(b_a, b_\beta) d\theta \ ,$$

and on the other hand, since $\mathbf{A}$ commutes with M_ϕ,

$$(\mathbf{A}(\phi b_\alpha), \mathbf{A} b_\beta) = (\phi \mathbf{A} b_\alpha, \mathbf{A} b_\beta) = \frac{1}{2\pi} \int_0^{2\pi} \phi(\theta)(g_\alpha(\theta), g_\beta(\theta)) d\theta .$$

Therefore $(g_\alpha(\theta), g_\beta(\theta)) = (b_\alpha, b_\beta)$ for θ outside a set $N_{\alpha\beta}$ of measure zero. Since $\mathcal{K}$ is separable $N = \cup N_{\alpha\beta}$ has measure zero, and

$$(A(\theta)b_\alpha, A(\theta)b_\beta) = (b_\alpha, b_\beta) \quad \text{for all } \alpha, \beta \text{ and all } \theta \notin N.$$

For an arbitrary element $x = \Sigma\, c_\alpha b_\alpha$ of $\mathcal{K}$, define $A(\theta)x = \Sigma\, c_\alpha A(\theta) b_\alpha$. This series converges for each $\theta \notin N$, and defines an isometry by the previous equation.

Now $\mathbf{A} b_\alpha = \hat{A} b_\alpha$ for each α; since both operators are linear and bounded, they agree on $\mathcal{K}$. The Fourier expansion of $f \in L^2(\mathcal{K})$ may be written as $\Sigma_{-\infty}^{+\infty}\, U^n x_n$ with $x_n \in \mathcal{K}$; since both operators commute with U, this gives $\mathbf{A} f = \hat{A} f$.

Suppose $\hat{A} = \hat{B}$. Then $A(\theta)b_\alpha = B(\theta)b_\alpha$ except on a set N_α of measure zero. Let $N = \cup N_\alpha$, and let K be a set of measure zero outside of which both A and B are bounded. *If* $\theta \notin N \cup K$, then $A(\theta)$ and $B(\theta)$ agree on a dense submanifold of $\mathcal{K}$, and therefore on $\mathcal{K}$.

COROLLARY. A subspace $\mathfrak{M}$ of $L^2(\mathcal{K})$ reduces U if and only if there is a measurable operator-valued function P such that $P(\theta)$ is a projection a.e. and such that

$$\mathfrak{M} = \{ f \mid f(\theta) \in \mathfrak{M}(\theta) \text{ a.e.} \},$$

where $\mathfrak{M}(\theta) = P(\theta)\mathcal{K}$.

Proof: If $\mathfrak{M}$ reduces U and $\mathbf{P}$ is the projection on $\mathfrak{M}$, then $\mathbf{P} = \hat{P}$ by the theorem. In the previous proof it was shown that $A(\theta)$ is an isometry a.e., and a similar argument shows that $P(\theta)$ is a projection a.e. Finally, the following are equivalent: $f \in \mathfrak{M}$, $\mathbf{P}f = f$, $(\mathbf{P}f)(\theta) = f(\theta)$ a.e., $P(\theta)f(\theta) = f(\theta)$ a.e., and $f(\theta) \in \mathfrak{M}(\theta)$ a.e.

REMARKS. 1) The theorem and its corollary reduce to §3.IV and its corollary when $\mathcal{K}$ is taken to be one-dimensional, i.e. the complex numbers.

2) One cannot simply define $A(\theta)$ by means of $A(\theta)g(\theta) = (\mathbf{A}g)(\theta)$. For this one would need, with the exception of a set of measure zero independent of g, that the value of $\mathbf{A}g$ at θ depends only on the value of g at θ.

3) The situation which results if $\mathcal{K}$ is allowed to be inseparable is evidently not well understood. [55].

4) A similar theorem is valid for any normal operator N. The circle is replaced by $\sigma(N)$ and Lebesgue measure by a measure determined by the spectral family of N and a family of vectors cyclic for the commutant of N. ([36, Ch. V and VIII]; [17])

DEFINITION. An essentially bounded measurable operator-valued function A is called *analytic* if $\hat{A}H^2(\mathcal{K}) \subset H^2(\mathcal{K})$. In this case the restriction $\hat{A} \mid H^2(\mathcal{K})$ is denoted by $\hat{A}_+$.

V. **Theorem.** An operator $\mathbf{A}$ on $H^2(\mathcal{K})$ commutes with the unilateral shift U_+ if and only if $\mathbf{A} = \hat{A}_+$ for some analytic operator-valued function A.

Proof: Extend $\mathbf{A}$ to $L^2(\mathcal{K})$ by setting $\mathbf{A}_1 U^n x = U^n \mathbf{A} x$ for $n < 0$ and $x \in \mathcal{K}$. Then $\mathbf{A}_1$ commutes with U, so $\mathbf{A}_1 = \hat{A}$ by IV. Therefore A is analytic and $\mathbf{A} = \hat{A}_+$.

Exercise. A is analytic if and only if $\hat{A}\mathcal{K} \subset H^2(\mathcal{K})$.

VI. Lemma. Let A be an analytic operator-valued function. Then the following statements are equivalent:

1. $\hat{A}_+$ and U_+^* commute;
2. $\hat{A}_+\mathcal{K} \subset \mathcal{K}$; and
3. A is constant a.e.

Proof: If $\hat{A}_+$ and U_+^* commute, then $U_+^* \hat{A}_+ \mathcal{K} = \hat{A}_+ U_+^* \mathcal{K} = \{0\}$, so $\hat{A}_+\mathcal{K}$ is contained in $\mathcal{K}$, the null space of U_+^*. Conversely, suppose $\hat{A}_+\mathcal{K} \subset \mathcal{K}$. Then $\hat{A}_+ U_+^* = 0 = U_+^* \hat{A}_+$ on $\mathcal{K}$, while for $x \in \mathcal{K}$ and $n > 0$,

$$\hat{A}_+ U_+^* U_+^n x = \hat{A}_+ U_+^{n-1} x = U_+^* U_+ \hat{A}_+ U_+^{n-1} x = U_+^* \hat{A}_+ U_+^n x,$$

so $\hat{A}_+$ and U_+^* commute on $U_+^n\mathcal{K}$.

Suppose $\hat{A}_+\mathcal{K} \subset \mathcal{K}$, and define an operator C on $\mathcal{K}$ by $C = \hat{A}_+ | \mathcal{K}$. If $\{b_a\}$ is an orthonormal basis of $\mathcal{K}$, then for each a there is a set N_a of measure zero such that $A(\theta)b_a = Cb_a$, $\theta \notin N_a$. Since $\mathcal{K}$ is separable $N = \cup N_a$ is of measure zero, and $A(\theta) = C$ for $\theta \notin N$. Conversely, if $A(\theta) = C$ a.e., then for all $x \in \mathcal{K}$, $(\hat{A}x)(\theta) = A(\theta)x = Cx$ a.e. so that $\hat{A}x \in \mathcal{K}$.

COROLLARY. The subspace $\mathfrak{M}$ of $H^2(\mathcal{K})$ reduces U_+ if and only if $\mathfrak{M} = H^2(\mathfrak{L})$ for some subspace $\mathfrak{L}$ of $\mathcal{K}$.

Proof: Let $\mathbf{P}$ be the projection on a reducing subspace $\mathfrak{M}$, so that $\mathbf{P}\mathcal{K} \subset \mathcal{K}$ by V and VI. This implies that $\mathfrak{L} = \mathbf{P}\mathcal{K}$ is closed. Since $\mathfrak{L} \subset \mathfrak{M}$ and $\mathfrak{M}$ is invariant it follows that $H^2(\mathfrak{L}) \subset \mathfrak{M}$. On the other hand, if $\Sigma_0^\infty U^n x_n \in \mathfrak{M}$, then

$$\Sigma U^n x_n = \mathbf{P}(\Sigma U^n x_n) = \Sigma U^n \mathbf{P} x_n \in H^2(\mathfrak{L}) .$$

The invariant subspace of the bilateral shift will be analyzed in the next two lemmas.

VII. Lemma. Each invariant subspace $\mathfrak{M}$ of the bilateral shift admits a unique decomposition $\mathfrak{M} = \mathfrak{M}_\infty \oplus \mathfrak{N}$ such that $\mathfrak{M}_\infty$ is reducing, and $\mathfrak{N}$ is invariant with $\cap_{n=0}^\infty U^n\mathfrak{N} = \{0\}$.

Proof: Let $\mathfrak{M}_\infty = \cap_0^\infty U^n\mathfrak{M}$ and $\mathfrak{N} = \mathfrak{M} \ominus \mathfrak{M}_\infty$. Then $U\mathfrak{M}_\infty = \mathfrak{M}_\infty$ so $\mathfrak{M}_\infty$ reduces. This implies the invariance of $\mathfrak{N}$, while $\cap U^n\mathfrak{N} = \{0\}$ is obvious, as is the uniqueness.

Since $\mathfrak{M}_\infty$ is taken care of by the corollary to IV, the next lemma deals with the summand $\mathfrak{N}$. First recall that an operator V which is isometric on a subspace $\mathfrak{I}$ of a Hilbert space $\mathfrak{H}$, and zero on the orthogonal complement, is called a *partial isometry*. The subspace $\mathfrak{I}$ is called the *initial space* of V, and $\mathfrak{F} = V\mathfrak{H}$ the *final space*. The projections on the initial and final spaces are given by V^*V and VV^* respectively. The following statements are easily seen to be equivalent: V is a partial isometry, V^*V is a projection, VV^* is a projection, and $VV^*V = V$.

VIII. Lemma. If $\mathfrak{N}$ is an invariant subspace of the bilateral shift U such that $\cap_{n=0}^\infty U^n\mathfrak{N} = \{0\}$, then there are a subspace $\mathfrak{L}$ of $\mathfrak{K}$ and a partial isometry $\mathbf{V}$ on $L^2(\mathfrak{K})$, such that the initial space of $\mathbf{V}$ is $L^2(\mathfrak{L})$, $\mathbf{V}$ commutes with U, and $\mathfrak{N} = \mathbf{V}H^2(\mathfrak{L})$.

Proof: The hypotheses and §1.I imply that $U \mid \mathfrak{N}$ is a unilateral shift, so that $\mathfrak{N} = \Sigma_0^\infty \oplus U^n \mathfrak{L}_0$, where $\mathfrak{L}_0 = \mathfrak{N} \ominus U\mathfrak{N}$. Since $\mathfrak{L}_0$ is wandering for U, a lemma of §1 gives $\dim \mathfrak{L}_0 \leq \dim \mathfrak{K}$. Let $\mathfrak{L}$ be a subspace of $\mathfrak{K}$ with the dimension of $\mathfrak{L}_0$ and W an isometry from $\mathfrak{L}$ onto $\mathfrak{L}_0$. Put $\mathbf{V} = 0$ on $L^2(\mathfrak{L})^\perp$ and

$$\mathbf{V}\left(\sum_{-\infty}^{\infty} U^n x_n\right) = \sum_{-\infty}^{\infty} U^n W x_n$$

if $x_n \in \mathfrak{L}$ for all n. Then $\mathbf{V}$ is a partial isometry that commutes with U, and

$$\mathbf{V}H^2(\mathfrak{L}) = \sum_0^\infty \oplus U^n W \mathfrak{L} = \sum_0^\infty \oplus U^n \mathfrak{L}_0 = \mathfrak{N} .$$

REMARK. If $\dim \mathfrak{L}_0 = \dim \mathfrak{K}$, $\mathbf{V}$ can be taken to be an isometry.

The situation regarding uniqueness is as follows. If $\mathfrak{L}'$, $\mathbf{V}'$ is another such pair, then $\mathbf{V}'H^2(\mathfrak{L}') = \mathbf{V}H^2(\mathfrak{L})$ implies $\mathbf{V}'\mathfrak{L}' = \mathbf{V}\mathfrak{L}$, and consequently there is a partial isometry W_0 on $\mathfrak{K}$ with initial space $\mathfrak{L}'$ and final space $\mathfrak{L}$ such that $\mathbf{V}' = \mathbf{V}W_0$ on $\mathfrak{K}$. If $\mathbf{W}$ is defined on $L^2(\mathfrak{K})$ by $\mathbf{W}(\Sigma\, U^n x_n) = \Sigma\, U^n W_0 x_n$, it follows readily that $\mathbf{V}' = \mathbf{V}\mathbf{W}$ and that $\mathbf{W}$ commutes with U. In fact $\mathbf{W} = \hat{W}$, where $W(\theta) = W_0$ for all θ. Summing up:

IX. Theorem. Each invariant subspace of the bilateral shift U on $L^2(\mathfrak{K})$ is of the form $\mathfrak{M}_\infty \oplus \mathbf{V}H^2(\mathfrak{L})$, where $\mathfrak{M}_\infty$ reduces U, $\mathfrak{L}$ is a subspace of $\mathfrak{K}$, and $\mathbf{V}$ is a partial isometry that commutes with U and has initial space $L^2(\mathfrak{L})$. The space

$\mathfrak{M}_\infty$ is unique. If $\mathfrak{L}'$ is a subspace of $\mathfrak{K}$ and $\mathbf{V}'$ is a partial isometry such that $\mathbf{V}'$ commutes with U, the initial space of $\mathbf{V}'$ is $L^2(\mathfrak{L}')$, and $\mathbf{V}'H^2(\mathfrak{L}') = \mathbf{V}H^2(\mathfrak{L})$, then $\mathbf{V}' = \mathbf{V}\mathbf{W}$ for some constant partial isometry $\mathbf{W}$.

The statement that $\mathbf{W}$ is constant means that it commutes with U, and that the corresponding operator-valued function is constant a.e. If $\mathfrak{K}$ is one-dimensional, this result reduces easily to VI of the previous section. The two summands do not occur simultaneously, and $\mathbf{V}$ is either zero or multiplication by a function of modulus one a.e.

Exercise. Let $\mathbf{V}$ be as in the theorem, so that $\mathbf{V} = \hat{V}$ for some operator-valued function V. Show that, almost everywhere, $V(\theta)$ is a partial isometry with initial space $\mathfrak{L}$.

X. **Theorem.** (Lax [31]; Halmos [22]). The subspace $\mathfrak{M}$ of $H^2(\mathfrak{K})$ is invariant for the unilateral shift U_+ if and only if it is the range of a partial isometry $\mathbf{V}$ that commutes with U_+. The partial isometry $\mathbf{V}$ is determined by $\mathfrak{M}$ up to right-multiplication by a constant partial isometry.

Proof: If $\mathfrak{M}$ is invariant for U_+, then viewed as a subspace of $L^2(\mathfrak{K})$, it is invariant for U, and so IX applies. A non-zero reducing subspace of U cannot lie in $H^2(\mathfrak{K})$, so that $\mathfrak{M}_\infty = \{0\}$ and $\mathfrak{M} = \mathbf{V}_1 H^2(\mathfrak{L})$ for some partial isometry $\mathbf{V}_1$ that commutes with U and has $L^2(\mathfrak{L})$ as initial space. It follows that $H^2(\mathfrak{K})$ is invariant for $\mathbf{V}_1$ and that $\mathbf{V} = \mathbf{V}_1 | H^2(\mathfrak{K})$ is as required.

The uniqueness statement follows in much the same way as that of IX, except that the form of the initial space of $\mathbf{V}$

is not specified in X. This is unnecessary since a partial isometry $\mathbf{V}$ in $H^2(\mathcal{K})$ that commutes with U_+ must have initial space of the form $H^2(\mathcal{L})$ for some $\mathcal{L} \subset \mathcal{K}$. This is shown by the following exercise, together with the corollary to VI.

Exercise. (R. G. Douglas) If U is an isometry and V a partial isometry that commutes with U, the the initial space of V reduces U.

Uniqueness can also be deduced from the following general result.

XI. Theorem. If $\mathbf{V}_1$ and $\mathbf{V}_2$ are partial isometries in $H^2(\mathcal{K})$ each commuting with U_+, then range $\mathbf{V}_1 \subset$ range $\mathbf{V}_2$ if and only if $\mathbf{V}_1 = \mathbf{V}_2\mathbf{W}$ for some partial isometry $\mathbf{W}$ that commutes with U_+. Equality holds if and only if $\mathbf{W}$ is constant.

Proof: The converse is trivial, so suppose that range $\mathbf{V}_1 \subset$ range $\mathbf{V}_2$. Then it is easy to see that $\mathbf{W} = \mathbf{V}_2^*\mathbf{V}_1$ is a partial isometry with $\mathbf{V}_1 = \mathbf{V}_2\mathbf{W}$. To show that $\mathbf{W}$ and U_+ commute, observe first that $\mathbf{V}_2(U_+\mathbf{W} - \mathbf{W}U_+) = 0$ since U_+ commutes with $\mathbf{V}_2$ and $\mathbf{V}_2\mathbf{W}$. On the other hand, range $\mathbf{W}$ is a subspace of the initial space of $\mathbf{V}_2$, and the latter is invariant for U_+, so that the range of $U_+\mathbf{W} - \mathbf{W}U_+$ is contained in the initial space of $\mathbf{V}_2$. These two statements imply that $U_+\mathbf{W} - \mathbf{W}U_+ = 0$.

If equality holds, the reverse inclusion implies in the same way that $\mathbf{W}^*$ commutes with U_+, and then $\mathbf{W}$ is constant by VI.

This theorem permits an important reformulation of the invariant subspace problem. The first reformulation was to show that each invariant subspace of U_+^* of dimension at least two contained a further invariant subspace. Taking orthogonal complements and keeping XI and V in mind, the problem becomes one of finding non-constant factorizations for analytic partial-isometry valued functions. In this form the problem has been thoroughly analyzed by Potapov [39], for the case in which $\mathcal{K}$ is finite-dimensional. For an interesting discussion, see [28, Lecture VIII].

For the invariant subspace problem, it is actually sufficient to consider analytic *unitary-valued* functions. To see this, recall that if T is a proper contraction on $\mathcal{K}$, then

$$\mathfrak{M} = \left\{ \sum_0^\infty U^n T^n x \mid x \in \mathcal{K} \right\}$$

is the subspace of $H^2(\mathcal{K})$ introduced by Rota (§2). It is invariant for U_+^*, so the orthogonal complement $\mathfrak{N} = H^2(\mathcal{K}) \ominus \mathfrak{M}$ is invariant for U_+ (and for U, if $\mathfrak{N}$ is regarded as a subspace of $L^2(\mathcal{K})$). A look at the proof of VIII shows that if

$$\mathfrak{L}_0 = \mathfrak{N} \ominus U\mathfrak{N}$$

is a *complete* wandering subspace for U, then the partial isometry $\mathbf{V}$ constructed there can be taken to be unitary. An equivalent condition is that the smallest subspace containing $\mathfrak{N}$ and reducing U is all of $L^2(\mathcal{K})$. Such an invariant subspace is said to have *full range*.

Notice first that $T^*x - Ux \in \mathfrak{N}$ for all $x \in \mathcal{K}$. Since

$$T^*x - Ux \in H^2(\mathcal{K}) ,$$

it must be shown that $T^*x - Ux \perp \mathfrak{M}$. If $y \in \mathcal{K}$,

$$(U^nT^ny, T^*x) = \frac{1}{2\pi}\int_0^{2\pi}(e^{in\theta}T^ny, T^*x)d\theta$$

$$\equiv \frac{1}{2\pi}\int_0^{2\pi}(e^{in\theta}T^{n+1}y, x)d\theta$$

$$= (U^nT^{n+1}y, x),$$

and therefore

$$\left(\sum_0^\infty U^nT^ny, T^*x - Ux\right) = \sum_0^\infty (U^nT^ny, T^*x) - \sum_0^\infty (U^nT^ny, Ux)$$

$$= \sum_1^\infty (U^nT^{n+1}y, x) - \sum_1^\infty (U^{n-1}T^ny, x) + (Ty, x) - (y, Ux)$$

$$= 0 .$$

Now suppose that f is orthogonal to $\Sigma_{-\infty}^{\infty} \oplus U^n\mathfrak{L}_0$. Then $U^nf \perp T^*x - Ux$ for all $x \in \mathcal{K}$ and all integers n, so that

$$0 = (T^*x - Ux, U^nf) = \frac{1}{2\pi}\int_0^{2\pi} e^{-in\theta}(T^*x - e^{i\theta}x, f(\theta)d\theta,$$

$$(T^*x - e^{i\theta}x, f(\theta)) = 0 \text{ a.e.},$$

$$(T - e^{-i\theta}I)f(\theta) = 0 \text{ a.e.},$$

and

$$f(\theta) = 0 \text{ a.e.},$$

since T is a proper contraction.

SECTION 5.

ANALYTIC REPRESENTATION OF THE UNILATERAL SHIFT

In this section a brief alternative treatment of some of the results of the previous sections will be given. This point of view has been developed by de Branges and Rovnyak [10, 11].

Let $\mathcal{K}$ be a Hilbert space. The space $\ell^2_+(\mathcal{K})$ is naturally identified with the space $\mathcal{K}(z)$ of formal power series $\Sigma_{n=0}^{\infty} z^n x_n$ with vector coefficients $x_0, x_1, \ldots \epsilon \mathcal{K}$ such that $\Sigma \|x_n\|^2 < \infty$. In $\mathcal{K}(z)$ the unilateral shift U_+ is "multiplication by z." Let D be the open unit disc in the complex plane.

I. If $f(z) = \Sigma z^n x_n$ is in $\mathcal{K}(z)$, then for each $w \epsilon D$, $\Sigma w^n x_n$ converges to an element $f(w)$ of $\mathcal{K}$.

Proof: For all integers k and m with $k \leq m$,

$$\|\sum_{n=k}^{m} w^n x_n\|^2 \leq (\Sigma |w|^n \|x_n\|)^2 \leq (\Sigma |w|^{2n})(\Sigma \|x_n\|^2)$$

$$\leq (1 - |w|^2)^{-1} \sum_{n=k}^{m} \|x_n\|^2 \to 0$$

as $k, m \to \infty$.

COROLLARY. $\|f(w)\| \leq (1 - |w|^2)^{-\frac{1}{2}} \|f(z)\|$.

Exercise. If P is the projection on $\mathcal{K}$, considered as a subspace of $\mathcal{K}(z)$ in the obvious way, then

$$f(w) = P(I - wU_+^*)^{-1} f(z) \ .$$

If B_0, $B_1, \ldots.$ are operators on $\mathcal{K}$, the formal power series $B(z) = \Sigma\, z^n B_n$ with operator coefficients acts formally on power series $f(z) = \Sigma\, z^n x_n$ with vector coefficients by the usual rule for multiplication of series: $B(z)f(z) = \Sigma\, z^n y_n$, where $y_n = \Sigma_{k=0}^{n} B_k x_{n-k}$ for $n \geq 0$.

II. $B(z)$ is a bounded operator on $\mathcal{K}(z)$ of norm at most M if and only if for every $w \in D$ the series $B(w) = \Sigma\, w^n B_n$ converges in the operator norm to a bounded operator on $\mathcal{K}$ of norm at most M.

Proof: Suppose $\|B(z)f(z)\| \leq M\,\|f(z)\|$ for all $f(z) \in \mathcal{K}(z)$. Then $M^2\|x\|^2 \geq \|B(z)x\|^2 = \Sigma\,\|B_n x\|^2$ for all $x \in \mathcal{K}$, so $\|B_n\| \leq M$ for $n \geq 0$, and consequently $\Sigma\, w^n B_n$ converges in the operator norm to a bounded operator $B(w)$ on $\mathcal{K}$ for all $w \in D$. Now

$$(f(w), x)_{\mathcal{K}} = (P(I - wU_+^*)^{-1} f(z), x)_{\mathcal{K}} = (f(z), (I - \bar{w}\, z)^{-1} x)_{\mathcal{K}(z)} \ .$$

Substituting $f(z) = B(z)(I - \bar{w}\, z)^{-1} y$ with $y \in \mathcal{K}$ gives

$$(B(w)y, x) = (1 - |w|^2)(B(z)(I - \bar{w}\, z)^{-1} y,\ (I - \bar{w}\, z)^{-1} x) \ ,$$

and therefore

$$|(B(w)y, x)|^2 \leq (1 - |w|^2)^2 M^2 \|(I - \bar{w}\, z)^{-1} y\|^2 \|(I - \bar{w}\, z)^{-1} x\|^2 \ .$$

Since $\|(I - \bar{w}\, z)^{-1} y\|^2$ is easily computed to be $(1 - |w|^2)^{-1}\|y\|^2$, it follows that $\|B(w)\| \leq M$.

Conversely, suppose that $\|B(w)\| \leq M$ for all $w \in D$. Let $f(z) = \Sigma\, z^n x_n$ be an element of $\mathcal{K}(z)$ and $B(z)f(z) = \Sigma\, z^n y_n$. By hypothesis $\|\Sigma\, w^n y_n\|^2 \leq M^2 \|\Sigma\, w^n x_n\|^2$ for all $w \in D$. Letting $w = re^{i\theta}$ and integrating this relation over θ gives $\Sigma\, r^{2n}\|y_n\|^2 \leq M^2\, \Sigma\, r^{2n}\|x_n\|^2$ for $0 \leq r < 1$. Therefore $\Sigma\, \|y_n\|^2 \leq M^2\, \Sigma\, \|x_n\|^2$, or equivalently,

$$\|B(z)f(z)\|^2 \leq M^2 \|f(z)\|^2 .$$

Exercise. $B(w) = P(I - wU_+^*)^{-1}B(z)|\mathcal{K}$.

III. An operator B on $\mathcal{K}(z)$ is of the form $B(z)$ if and only if it commutes with the unilateral shift.

Proof: Suppose B commutes with the shift. For $x \in \mathcal{K}$ let $Bx = \Sigma\, z^n x_n$, and define $B_n x = x_n$, so that $Bx = \Sigma\, z^n B_n x$. These operators are linear and bounded by $\|B\|$, and if $f(z) = \Sigma\, z^n x_n$ is an element of $\mathcal{K}(z)$, then

$$Bf(z) = \sum_{k=0}^{\infty} z^k B x_k = \sum_{k=0}^{\infty} z^k \left(\sum_{n=0}^{\infty} z^n B_n x_k \right) = B(z)f(z)$$

since B commutes with the shift.

SECTION 6. CONTRACTIONS

In section 2 it was shown that any contraction whose powers tend strongly to zero is part of a backward shift. The backward shift is a coisometry (the adjoint of an isometry), so it is natural to ask what operators can be obtained as parts of general coisometries. It will be shown that all contractions arise in this way, that the coisometry corresponding to a given contraction T may be taken to be minimal in a certain sense, and that then it is unique up to unitary equivalence. This coisometry C_T will be referred to as *the minimal coisometric extension of T*. Recall that an invariant subspace for an operator is said to be full or to have full range if the smallest reducing subspace that contains it is the whole space. The result then goes as follows:

I. Theorem. If T is a contraction on a Hilbert space $\mathcal{H}$, there are a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ and a coisometry C on $\mathcal{K}$ such that $\mathcal{H}$ is a full invariant subspace for C and $C \mid \mathcal{H} = T$. If $(\mathcal{L}, D)$ is another such pair there is an isometry J from $\mathcal{K}$ onto $\mathcal{L}$ such that $DJ = JC$ and J is the identity on $\mathcal{H}$.

Proof: Let $S = (I - TT^*)^{1/2}$ and let $\mathcal{S}$ be the closure of the range $S\mathcal{H}$ of S. The operator matrix (top, p. 50) defines an operator C on the space $\mathcal{K} = \mathcal{H} \oplus \mathcal{S} \oplus \mathcal{S} \oplus \cdots$ (elements of $\mathcal{K}$ being treated as column vectors). This operator satisfies

$$\begin{bmatrix} T & S & 0 & 0 & 0 & \\ 0 & 0 & I & 0 & 0 & \\ 0 & 0 & 0 & I & 0 & \\ 0 & 0 & 0 & 0 & I & \\ \vdots & & & & & \ddots \end{bmatrix}$$

$CC^* = I$ and consequently is a coisometry. The space $\mathcal{H}$, identified with the subspace $\mathcal{H} \oplus \{0\} \oplus \{0\} \oplus \cdots$, is obviously invariant with the property $C \mid \mathcal{H} = T$. The computations

$$C^*(x \oplus 0 \oplus 0 \oplus \cdots) = T^*x \oplus Sx \oplus 0 \oplus \cdots$$

$$C^*(0 \oplus y \oplus 0 \oplus \cdots) = 0 \oplus 0 \oplus y \oplus 0 \oplus \cdots$$

$$C^*(0 \oplus 0 \oplus y \oplus 0 \oplus \cdots) = 0 \oplus 0 \oplus 0 \oplus y \oplus 0 \oplus \cdots$$

and so forth, for $x \in \mathcal{H}$ and $y \in \mathcal{S}$, show that $\mathcal{H}$ is full. It also follows that the elements $\Sigma_{k=0}^{n} C^{*k}x_k$, for all $n \geq 0$ and $x_0, x_1, \ldots, x_n \in \mathcal{H}$, are dense in $\mathcal{K}$. Define an operator J on this dense submanifold of $\mathcal{K}$ to the corresponding dense submanifold of $\mathcal{L}$ by

$$J\left(\sum_{k=0}^{n} C^{*k}x_k\right) = \sum_{k=0}^{n} D^{*k}x_k \quad .$$

Then J is unambiguously defined and isometric, for

$$\begin{aligned} \|\Sigma\, D^{*k}x_k\|^2 &= \sum_{k,m} (D^{*k}x_k, D^{*m}x_m) \\ &= \sum_{k \leq m} (D^{m-k}x_k, x_m) + \sum_{k > m} (x_k, D^{k-m}x_m) \end{aligned}$$

$$= \sum_{k \leq m} (T^{m-k}x_k, x_m) + \sum_{k > m} (x_k, T^{k-m}x_m)$$

$$= \|\Sigma\, C^{*k}x_k\|^2 \quad .$$

Therefore J has a unique isometric extension from $\mathcal{K}$ onto $\mathcal{L}$. It is obvious that J is the identity on $\mathcal{H}$ and that $JC^* = D^*J$. Multiplication of the last equation by J^* on the right and left gives $C^*J^* = J^*D^*$, and so $DJ = JC$ as required.

It should be noticed that the uniqueness result, together with §2.II, implies that the isometry C_T^* is pure (i.e., a unilateral shift) if and only if $T^n \to 0$ strongly.

A proof of this theorem along the lines of the special case §2.II has been given by R. G. Douglas [15], and goes as follows. If, as in the proof of §2.II, $R = (I - T^*T)^{1/2}$, $\mathcal{R} = (R\mathcal{H})^-$, and $W: \mathcal{H} \to H^2(\mathcal{R})$ is defined by

$$Wx = (Rx, RTx, RT^2x, \ldots) \quad ,$$

then $\|Wx\|^2 = \|x\|^2 - \lim \|T^n x\|^2$. Now $\{T^{*n}T^n\}$ is a decreasing sequence of positive contractions, which therefore converges strongly to a positive contraction. If A is the positive square root of this contraction, then

$$\|Ax\|^2 = (A^2x, x) = \lim \|T^n x\|^2 \quad ,$$

and therefore $\|Wx\|^2 + \|Ax\|^2 = \|x\|^2$. With $\mathcal{G} = (A\mathcal{H})^-$, this means that $V: x \to Wx \oplus Ax$ is an isometry from $\mathcal{H}$ into $H^2(\mathcal{R}) \oplus \mathcal{G}$. The mapping $Ax \to ATx$ is readily seen to define an isometry G on $\mathcal{G}$ such that $GA = AT$. Since $WT = U_+^* W$, it follows that $VT = (U_+^* \oplus G)V$, and therefore that T is unitarily equivalent to $(U_+^* \oplus G)|V\mathcal{H}$. The operator $U_+^* \oplus G$ need not be a coisometry, but this can be remedied by means of the extension theorem §1.III for isometries. If the isometry G is extended to a unitary operator F, then $U_+^* \oplus F$ is the desired coisometry.

The ideas in this proof are often more useful than those in the first one. For example, an analogous theorem for one-parameter semigroups of contractions may be proved in this way, with the help of some basic information about infinitesimal generators (cf. §9.III).

The positive contraction A constructed above embodies useful information about the contraction T. For example, it is not difficult to see that T is similar to an isometry if and only if A is invertible. It is also useful, as pointed out by R.G. Douglas, in proving the result, due to Nagy and Foias [52], that any contraction has a largest unitary summand. A contraction is called *completely nonunitary* if this summand vanishes.

II. Theorem. Let T be a contraction, and let A and A_* be the positive contractions defined by $A^2 = \text{str.}\lim T^{*n}T^n$ and $A_*^2 = \text{str.}\lim T^nT^{*n}$. Then $\mathfrak{U} = \{x | Ax = A_*x = x\}$ is the largest reducing subspace of T on which T is unitary.

Proof: It is obvious that $\mathfrak{U}$ is a closed subspace. Notice that $T^*A^2T = A^2$ and $TA_*^2T^* = A_*^2$. The following elementary facts will also be useful: for any contraction T and vector x, the statements $\|Tx\| = \|x\|$ and $T^*Tx = x$ are equivalent; if $T \geq 0$ they are equivalent to $Tx = x$.

Now $A^2 \leq T^*T$ implies $\|Ax\|^2 \leq \|Tx\|^2 \leq \|x\|^2$ for any vector x, so that if $x \in \mathfrak{U}$, then $\|Tx\| = \|x\|$ and $T^*Tx = x$.

In particular T is isometric on $\mathfrak{U}$, and in the same way, so is T^*. This will imply that T is unitary on $\mathfrak{U}$, once it is shown that $\mathfrak{U}$ reduces. For the latter, if $x \in \mathfrak{U}$ then

$$\|ATx\|^2 = (T^*A^2Tx, x) = (A^2x, x)$$
$$= \|x\|^2 \geq \|Tx\|^2 \geq \|ATx\|^2,$$

and so $\|ATx\| = \|Tx\|$. Therefore $ATx = Tx$ by the remark above, since $0 \leq A \leq I$. The same inequalities give $\|Tx\| = \|x\|$, so that $T^*Tx = x$ and

$$A_*^2Tx = TA_*^2T^*Tx = TA_*^2 = Tx\,,$$

and therefore $A_*Tx = Tx$, again using the above remark. Consequently $\mathfrak{U}$ is invariant for T. In the same way $\mathfrak{U}$ is invariant for T^*, which proves that $\mathfrak{U}$ reduces T.

If the vector x is a member of a subspace which reduces T and in which T is unitary, then $\|T^nx\| = \|x\|$ for all $n \geq 0$, and so $\|Ax\|^2 = \lim \|T^nx\|^2 = \|x\|^2$. Then $A^2x = x$ by the remark above, which implies $Ax = x$ since A is non-negative. Dually, $A_*x = x$, and consequently $x \in \mathfrak{U}$.

This decomposition is often useful in reducing questions about contractions to the case of completely nonunitary contractions. Another useful decomposition, due to Foguel [18], runs as follows.

III. Theorem. Let T be a contraction, and let

$$\mathfrak{Z}(T) = \{x \mid T^nx \to 0 \text{ weakly}\}\,.$$

Then $\mathfrak{Z}(T) = \mathfrak{Z}(T^*)$, $\mathfrak{Z}(T)$ is a reducing subspace, and T is unitary on $\mathfrak{Z}(T)^{\perp}$.

Proof: The first conclusion follows from the equivalence of the conditions $(T^n x, x) \to 0$ and $T^n x \to 0$ weakly. In one direction this is trivial; for the other, let C be the minimal coisometric extension of T^* on $\mathcal{K} \supset \mathcal{H}$. Then since

$$(C^{*n}x, x) = (x, C^n x) = (x, T^{*n}x) = (T^n x, x)$$

$(T^n x, x) \to 0$ implies $(C^{*n}x, x) \to 0$. Now $C^k C^{*n} = C^{*(n-k)}$ for $n \geq k$, and therefore $\lim (C^{*n}x, C^{*k}x) = 0$ for all $k \geq 0$. The relation $(C^{*n}x, y) \to 0$ now follows, first for y in the closed linear span of the vectors $C^{*k}x$ (since they are uniformly bounded), second for y orthogonal to this span, and consequently for all $y \in \mathcal{K}$. If $y \in \mathcal{H}$, then

$$(T^n x, y) = (x, T^{*n}y) = (x, C^n y) = (C^{*n}x, y) \to 0 .$$

Since $\mathcal{Z}(T)$ is a closed subspace invariant for T, the condition $\mathcal{Z}(T) = \mathcal{Z}(T^*)$ implies that $\mathcal{Z}(T)$ reduces. To see that T is unitary in $\mathcal{Z}(T)^{\perp}$, it will first be shown that $(I - T^*T)\mathcal{H} \subset \mathcal{Z}(T)$. For this,

$$\begin{aligned} |(T^{*n}(I - T^*T)x, y)|^2 &= |(x, (I - T^*T)T^n y)|^2 \\ &\leq \|x\|^2 \|T^n y - T^*T T^n y\|^2 \\ &\leq 2\|x\|^2(\|T^n y\|^2 - \mathrm{Re}(T^n y, T^*T T^n y)) \\ &= 2\|x\|^2(\|T^n y\|^2 - \|T^{n+1}y\|^2) , \end{aligned}$$

and the latter converges to zero, since

$$\sum_{n=0}^{\infty} (\|T^n y\|^2 - \|T^{n+1}y\|^2) = \|y\|^2 - \lim \|T^n y\|^2 < \infty .$$

Now suppose that x is orthogonal to $\mathcal{Z}(T)$. Then so is

$(I - T^*T)x$, since $\mathcal{Z}(T)$ reduces T, and therefore $(I - T^*T)x = 0$. This means that T is isometric on $\mathcal{Z}(T)^\perp$. In the same way so is T^*, and the proof is finished.

COROLLARY. If T is a completely nonunitary contraction, then $T^n \to 0$ weakly.

As to the connection between these two decompositions, it is of course true that $\mathcal{Z}(T)^\perp \subset \mathfrak{U}$. That equality need not hold can be seen by taking T to be a bilateral shift, or even a direct summand (restriction to a reducing subspace) of a bilateral shift, in which case both $\mathcal{Z}(T)$ and $\mathfrak{U}$ are the entire space. A unitary operator is said to be *absolutely continuous* if its spectral measure is absolutely continuous with respect to Lebesgue measure on the unit circle, and *singular* if its spectral measure is singular with respect to Lebesgue measure. It is not difficult to see that any unitary operator is uniquely the direct sum of an absolutely continuous unitary and a singular unitary, and that a unitary operator is absolutely continuous if and only if it is a direct summand of a bilateral shift. With these facts, the above remarks mean that $T|\mathcal{Z}(T)^\perp$ is singular.

J. P. Williams has pointed out that Foguel's decomposition permits an easy proof of the following result:

If $P_1, P_2, \ldots, P_k$ are projections, then $(P_1 P_2 \cdots P_k)^n$ converges weakly to $P_1 \wedge P_2 \wedge \cdots \wedge P_k$ (the projection on the intersection of the ranges of the P_i). For if $T = P_1 P_2 \cdots P_k$, then $T^n \to 0$ weakly on $\mathcal{Z}(T)$, and T is the identity on $\mathcal{Z}(T)^\perp$ (since it is isometric there). Thus T^n converges weakly to the projection on $\mathcal{Z}(T)^\perp = (P_1 \wedge P_2 \wedge \cdots \wedge P_k)\mathcal{H}$.

Halperin [26] shows that the convergence is actually strong.

Exercises. 1. The contraction T is an isometry if and only if $A = I$.

2. The contraction T is similar to an isometry if and only if A is invertible.

3. The null space of A is invariant for T.

4. If T is normal, then A is a projection.

5. If T is a contraction, then $\|T^n x\| = \|x\|$ if and only if $T^*Tx = x$.

6. A vector x is in $\mathfrak{U}$ if and only if $\|T^n x\| = \|x\|$ and $\|T^{*n} x\| \Omega \|x\|$ for all $n \geq 0$.

7. If U is a unitary operator, there is a unique reducing subspace $\mathfrak{M}$ such that $U \mid \mathfrak{M}$ is absolutely continuous and $U \mid \mathfrak{M}^\perp$ is singular.

8. A unitary operator is absolutely continuous if and only if it is a direct summand of a bilateral shift.

9. If C is a coisometry and $\mathfrak{M}$ a full invariant subspace, the isometry $C^* \mid \mathfrak{M}^\perp$ is pure.

SECTION 7.

DILATIONS

If T is an operator on a Hilbert space $\mathcal{H}$, and P is the projection on a subspace $\mathfrak{M}$, the operator $PT|\mathfrak{M}$ is called the *compression of* T to $\mathfrak{M}$, and T is called a *dilation of* $PT|\mathfrak{M}$. Recall that $(PT|\mathfrak{M})^* = PT^*|\mathfrak{M}$. If $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is the matrix of T with respect to the decomposition $\mathcal{H} = \mathfrak{M} \oplus \mathfrak{M}^{\perp}$, then $PT|\mathfrak{M} = A$.

It is obvious that any compression of a unitary operator is a contraction. Conversely:

I. Any contraction has a unitary dilation.

Proof: If T is a contraction on $\mathcal{H}$, then

$$W = \begin{bmatrix} T & S \\ R & -T^* \end{bmatrix}$$

is unitary on $\mathcal{H} \oplus \mathcal{H}$, where $R = (I - T^*T)^{1/2}$ and $S = (I - TT^*)^{1/2}$. This is verified by computation, with the help of the relation $TR = ST$. To see the latter, observe that $TR^2 = S^2T$, so that $Tp(R^2) = p(S^2)T$ for all polynomials p. Since the polynomials in x^2 are uniformly dense in $C[0,1]$, it follows that $TR = ST$.

In similar fashion, compressions of projections may be identified as the class of all positive contractions.

II. Any positive contraction can be dilated to a projection.

Proof: If A is a positive contraction on $\mathcal{H}$, the operator

$$\begin{bmatrix} A & (A-A^2)^{1/2} \\ (A-A^2)^{1/2} & I-A \end{bmatrix}$$

on $\mathcal{H} \oplus \mathcal{H}$ is a projection.

These results, due respectively to Halmos and Michael, are discussed in [24, Problem 177]. If W is a unitary dilation of $T = PW|\mathcal{H}$, it need not be true that $T^2 = PW^2|\mathcal{H}$. A unitary dilation W such that $T^n = PW^n|\mathcal{H}$ for all $n > 0$ will be called *strong*. Of course then $T^{*n} = PW^{*n}|\mathcal{H}$ for all $n > 0$. A result of the preceding section makes it easy to see that any contraction T has a strong unitary dilation. For this recall the two constructions in that section of the minimal coisometric extension C, taking place on the spaces $\mathcal{H} \oplus H^2(\mathcal{S})$ and $H^2(\mathcal{R}) \oplus \mathcal{F}$ (where $R = (I-T^*T)^{1/2}$, $S = (I-TT^*)^{1/2}$, $\mathcal{R} = (RH)^-$, $\mathcal{S} = (S\mathcal{H})^-$). From the uniqueness it follows that there is a decomposition

$$\mathcal{H} \oplus H^2(\mathcal{S}) = H^2(\mathcal{R}') \oplus \mathcal{F}' ,$$

where $\mathcal{R}'$ is isomorphic to $\mathcal{R}$ and $\mathcal{F}'$ to $\mathcal{F}$. and $C = U_+^* \oplus F'$, where U_+ is the unilateral shift on $H^2(\mathcal{R}')$ and F' is unitary on $\mathcal{F}'$. Then $W = U^* \oplus F'$ on $\mathcal{K} = L^2(\mathcal{R}') \oplus \mathcal{F}'$ is a strong unitary dilation of T. This is clear since the compression of W^n to $H^2(\mathcal{R}') \oplus \mathcal{F}'$ is C^n for all $n > 0$.

A uniqueness statement, similar to that for the coisometric extension and proved in the same way, is valid for strong unitary dilations. Call a dilation W on a space $\mathcal{K} \supset \mathcal{H}$ *minimal* if the smallest subspace containing $\mathcal{H}$ and reducing W is all of $\mathcal{K}$. The minimality of the dilation just constructed follows from that of C. Moreover, for any such dilation, the elements $\Sigma_{k=-n}^{n} W^k x_k$, $x_k \in \mathcal{H}$, form a dense submanifold, and this fact implies that the minimal strong unitary dilation is unique up to a unitary equivalence leaving the elements ments of $\mathcal{H}$ fixed.

III. Theorem. Any contraction T on a Hilbert space $\mathcal{H}$ has an essentially unique minimal strong unitary dilation W. The dilation space may be decomposed as

$$\mathcal{K} = \left(\sum_{n=1}^{\infty} \oplus W^n \mathfrak{R}' \right) \oplus \mathcal{H} \oplus \left(\sum_{n=1}^{\infty} \oplus W^{*n} \mathcal{S} \right),$$

where $\mathfrak{R}'$ and $\mathcal{S}$ are wandering subspaces for W, $\mathfrak{R}'$ is isomorphic to $((I - T^*T)^{1/2}\mathcal{H})^-$, and $\mathcal{S} = ((I - TT^*)^{1/2}\mathcal{H})^-$.

Proof: Only the last statement will be considered. This follows from the relations

$$\mathcal{K} = L^2(\mathfrak{R}') \ominus H^2(\mathfrak{R}')) \oplus \mathcal{H} \oplus H^2(\mathcal{S}),$$

$$L^2(\mathfrak{R}') \ominus H^2(\mathfrak{R}') = \sum_{n=1}^{\infty} \oplus W^n \mathfrak{R}',$$

and

$$H^2(\mathcal{S}) = \sum_{n=1}^{\infty} \oplus W^{*n} \mathcal{S} .$$

IV. $L^2(\mathfrak{R}')^\perp \cap L^2(\mathcal{S})^\perp$ is the largest subspace of $\mathcal{H}$ which reduces T and in which T is unitary.

Proof: Suppose $\mathfrak{N}$ reduces T and $T|\mathfrak{N}$ is unitary. Then $I - T^*T = 0$ on $\mathfrak{N}$, and so $\mathfrak{N} \subset \mathcal{F}' = L^2(\mathfrak{R}')^\perp$ from the construction of the coisometric extension. Dually $\mathfrak{N} \subset L^2(\mathcal{S})^\perp$. The space $L^2(\mathfrak{R}')^\perp \cap L^2(\mathcal{S})^\perp$ obviously reduces W, and by

III is a subspace of $\mathcal{H}$. It follows that $T = W$ on this space, and consequently T is unitary there.

V. The minimal strong unitary dilation of a completely nonunitary contraction is absolutely continuous.

Proof: Let $W = \int_0^{2\pi} e^{it} dE_t$ be the minimal dilation of T. The spaces $L^2(\mathfrak{R}')$ and $L^2(\mathcal{S})$ reduce W, and in each W is a bilateral shift. Therefore the measure $\|E(\cdot)z\|^2$ is absolutely continuous if z is a vector in either of these spaces. If $z = x + y$ is an element of $L^2(\mathfrak{R}') + L^2(\mathcal{S})$, then since $\|E(\cdot)z\|^2 \leq (\|E(\cdot)x\| + \|E(\cdot)y\|)^2$, $\|E(\cdot)z\|^2$ is absolutely continuous. But when T is completely nonunitary this manifold is dense by IV, and the result follows.

In many cases W is actually a bilateral shift.

VI. In each of the following situations, the minimal strong unitary dilation is a bilateral shift:

(i) $T^n \to 0$ strongly, or dually;

(ii) T is completely nonunitary and the rank of $I - T^*T$ is infinite, or dually.

Proof: In case (i) it has been observed previously that the minimal coisometric extension is pure, which implies that the minimal strong unitary dilation is a shift.

In case (ii) W has a direct summand which is a bilateral shift of infinite multiplicity by III. But W is absolutely continuous by V. By a well-known technique of multiplicity theory, the shift summand of W will "absorb" the rest of W (up to unitary equivalence), and so W is a bilateral shift.

The minimal dilation W of a completely nonunitary contraction T need not be a shift. The question of the spectral type of W when both $I - T^*T$ and $I - TT^*$ are of finite rank has been resolved by Nagy and Foias [53]. Those W that can occur are precisely as follows: the bilateral shift restricted to reducing subspaces of the form

$$L^2(\mu_1) \oplus \cdots \oplus L^2(\mu_n) \quad ,$$

where μ_i is Lebesgue measure on a subset S_i of the circle, $S_1 \supset S_2 \supset \cdots \supset S_n$, and at least half the S_i are the whole circle.

The results of this section are due mainly to Nagy and Foias, and can be found, along with many other things, in their series of papers on contractions. The present treatment, with its geometrical flavor, is essentially that of Douglas [15]. Other relevant papers are [20, 25, and 44].

It should be mentioned that the existence of the minimal strong unitary dilation can be established in the same way as that of the minimal coisometric extension, namely, by displaying a matrix for it. The matrix in question is

$$\begin{bmatrix} \ddots & & & & & & 0 \\ & I & & & & & \\ & & I & & & & \\ & & & R & -T^* & & \\ & & & T & S & & \\ & & & & & I & \\ & & & & & & I \\ 0 & & & & & & \ddots \end{bmatrix}$$

acting on $(\Sigma_{n<0} \otimes \mathfrak{R}) \otimes \mathcal{H} \oplus (\Sigma_{n>0} \oplus \mathcal{S})$, where $R = (I-T^*T)^{1/2}$,

$S = (I - TT^*)^{1/2}$, $\mathcal{R} = (R\mathcal{H})^-$, and $\mathcal{S} = (S\mathcal{H})^-$. Here the $(0,0)$ location of the matrix contains T.

Now several applications to von Neumann's theory of spectral sets will be given [38]. A closed subset X of the complex plane is called a *spectral set* for an operator T if it contains the spectrum of T, and if, for any rational function r with poles lying outside of X,

$$\|r(T)\| \leq \sup\{|r(z)| \mid z \in X\} .$$

VII. Theorem. The closed unit disc $\bar{D}$ is a spectral set for any contraction T.

Proof: Let f be holomorphic in a region containing D, and let W be a strong unitary dilation of T. If $f(z) = \Sigma_{n=0}^{\infty} a_n z^n$, then $\Sigma_{n=0}^{\infty} a_n T^n$ converges in norm to a bounded operator $f(T)$, and

$$f(T) = \sum a_n (PW^n \mid \mathcal{H}) = Pf(W) \mid \mathcal{H} .$$

Now from the spectral mapping theorem,

$$\sigma(f(W)) = f(\sigma(W)) \subset \{f(z) \mid z \in \bar{D}\} ,$$

and since the normality of $f(W)$ implies that its norm and spectral radius are equal, it follows that

$$\|f(T)\| \leq \|f(W)\| \leq \sup\{|f(z)| \mid z \in \bar{D}\} .$$

For the next application, recall that $\bar{W}(T)$ denotes the closure of the numerical range $\{(Tx, x) \mid \|x\| = 1\}$ of the operator T.

VIII. Theorem. (J. P. Williams [59].) If C is a closed convex set containing $\sigma(T)$ in its interior, there is an invertible operator S such that $\bar{W}(S^{-1}TS) \subset C$.

Proof: The following fact will be needed: If H is a convex spectral set of T, then $\bar{W}(T) \subset H$. It suffices to prove this for H a closed half-plane, and by translation and rotation it can be assumed that $H = \{z \mid \operatorname{Re} z \geq 0\}$. It must be shown that $\operatorname{Re}(Tx, x) \geq 0$ for all vectors x. Since $|(1-z)(1+z)^{-1}|$ is bounded by 1 on H, it follows from the definition of spectral set that $\|(I-T)(I+T)^{-1}\| \leq 1$. This is equivalent to $\|(I-T)x\| \leq \|(I+T)x\|$ for all x; squaring and expanding gives $\operatorname{Re}(Tx, x) \geq 0$ as required.

Now let V be the interior of C, let f be a conformal map of the open unit disc D onto V, and let g be the inverse map. Then D contains the spectrum of $g(T)$, and so by §2.1 there is an invertible operator S such that $\|S^{-1}g(T)S\| = r < 1$. Let D_1 be the closed disc of radius r. By VII D_1 is a spectral set of $g(S^{-1}TS) = S^{-1}g(T)S$. But f is a uniform limit of polynomials on D_1, from which it follows readily that $f(D_1)$ is a spectral set for $f(g(S^{-1}TS)) = S^{-1}TS$. Since C contains $f(D_1)$ it is also a spectral set, and because it is convex the assertion of the first paragraph gives $\bar{W}(S^{-1}TS) \subset C$.

SECTION 8

NAIMARK'S THEOREMS ON DILATIONS

A complex-valued function ϕ defined on a group G is called *positive-definite* if the inequality

$$\sum_{i,j=1}^{n} \phi(g_j^{-1} g_i)\lambda_i \overline{\lambda_j} \geq 0$$

holds for every choice of group elements $g_1, \ldots, g_n$ and complex numbers $\lambda_1, \ldots, \lambda_n$. It follows that

$$\phi(e) \geq 0 \text{ and } \phi(g^{-1}) = \overline{\phi(g)} ,$$

the former by taking $n = 1$, and the latter by taking $n = 2$, $g_1 = e$ (the identity), $g_2 = g$, $\lambda_1 = 1$, and $\lambda_2 = \lambda$.

A function $U: g \to U(g)$ assigning to each group element a unitary operator on a Hilbert space $\mathcal{K}$ is called a *unitary representation* of G if $U(e) = I$ and $U(gh) = U(g)U(h)$ for all $g, h \in G$.

These concepts are intimately related. If U is a unitary representation and x is a vector, a simple computation shows that the function $\phi(g) = (U(g)x, x)$ is positive-definite. On the other hand:

1. Theorem. If ϕ is a positive-definite function on a group G, there is a unitary representation U of G and a vector x such that $\phi(g) = (U(g)x, x)$ for all $g \in G$.

Proof: Let $\mathcal{K}_1$ be the complex vector space of all complex-valued functions ξ on G which vanish except on a finite subset. For $\xi, \eta \in \mathcal{K}_1$, define

$$(\xi,\eta)_1 = \sum_{g,h} \phi(h^{-1}g)\xi(g)\overline{\eta(h)} \ .$$

This is a bilinear functional which satisfies $(\xi,\xi)_1 \geq 0$ since ϕ is positive-definite. The relation $\phi(g^{-1}) = \overline{\phi(g)}$ evidently implies that $(\xi,\eta)_1 = \overline{(\eta,\xi)}_1$. Consequently the C.B.S. inequality

$$|(\xi,\eta)_1|^2 \leq (\xi,\xi)_1(\eta,\eta)_1$$

is valid. It follows that $\mathcal{K}_0 = \{\xi \,|\, (\xi,\xi)_1 = 0\}$ is a subspace of $\mathcal{K}_1$, and that

$$(\xi + \mathcal{K}_0, \eta + \mathcal{K}_0) = (\xi, \eta)_1$$

is a well-defined bilinear functional which makes the quotient space $\mathcal{K}_1/\mathcal{K}_0$ a pre-Hilbert space. Let $\mathcal{K}$ be the completion.

For $k \in G$ define a linear transformation U_k on $\mathcal{K}_1$ by $(U_k\xi)(g) = \xi(k^{-1}g)$. Then $(U_k\xi, U_k\eta)_1 = (\xi,\eta)_1$, so U_k leaves leaves $\mathcal{K}_0$ invariant and induces a linear isometry on $\mathcal{K}_1/\mathcal{K}_0$, which has a unique linear isometric extension $U(k)$ to $\mathcal{K}$. It is easy to see that $U(e) = I$ and $U(g)U(h) = U(gh)$ for all $g, h \in G$, and so U is a unitary representation. If ξ is defined by $\xi(e) = 1$ and $\xi(g) = 0$ for all $g \neq e$, then

$$(U(k)(\xi + \mathcal{K}_0), \xi + \mathcal{K}_0) = (U_k\xi,\xi)_1 = \sum \phi(h^{-1}g)\xi(k^{-1}g)\overline{\xi(h)}$$

$$= \phi(k) \ ,$$

and the proof is complete.

Let $\mathcal{H}$ be a complex Hilbert space. An operator-valued function $A\colon G \to \mathcal{B}(\mathcal{H})$ is called *positive-definite* if the inequality

$$\sum_{i,j=1}^{n} (A(g_j^{-1} g_i)x_i, x_j) \geq 0$$

holds for every choice of group elements $g_1, \ldots, g_n$ and vectors $x_1, \ldots, x_n$. If $\mathcal{H}$ is one-dimensional this concept reduces to the previous one, and as in that case

$$A(e) \geq 0 \text{ and } A(g^{-1}) = A(g)^*$$

for all $g \in G$. Again, much as before, such functions arise by compressing a unitary representation on a space $\mathcal{K}$ to a subspace $\mathcal{H}$, and conversely, such a function has a unitary dilation. Here $\mathcal{K}_1$ consists of all functions $\xi\colon G \to \mathcal{H}$ which vanish except on a finite subset, and

$$(\xi,\eta)_1 = \sum_{g,h} (A(h^{-1}g)\xi(g), \xi(h)) \ .$$

For each $x \in \mathcal{H}$ define $\xi_x \in \mathcal{K}_1$ by $\xi_x(e) = x$ and $\xi_x(g) = 0$ for $g \neq e$. Then $x \to \xi_x + \mathcal{K}_0$ is a linear transformation from $\mathcal{H}$ into $\mathcal{K}$, and because

$$(\xi_x + \mathcal{K}_0, \xi_y + \mathcal{K}_0) = (\xi_x, \xi_y)_1 = \Sigma (A(h^{-1}g)\xi_x(g), \xi_y(h))$$
$$= (A(e)x, y) \ ,$$

$\mathcal{H}$ may be considered to be a subspace of $\mathcal{K}$ provided that $A(e) = I$. If this identification is made, then

$$(U(k)x, y) = (U_k\xi_x, \xi_y)_1 = (A(k)x, y)$$

for all $x, y \in \mathcal{H}$. This is equivalent to $A(k) = PU(k)|\mathcal{H}$, where P is the projection of $\mathcal{K}$ onto $\mathcal{H}$. This proves:

II. Theorem. If G is a group, $\mathcal{H}$ is a complex Hilbert space, and $A: G \to \mathcal{B}(H)$ is positive-definite with $A(e) = I$, there is a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a unitary representation U of G on $\mathcal{K}$ such that $A = PU \mid \mathcal{H}$.

It is not difficult to prove a uniqueness statement similar to those of the preceding section, namely: the smallest subspace containing $\mathcal{H}$ and reducing U is all of $\mathcal{K}$; if V is any unitary dilation of A on a space $\mathcal{L}$ which is minimal in this sense, there is an isomorphism W of $\mathcal{K}$ onto $\mathcal{L}$ such that $V(g)W = WU(g)$ for all $g \in G$ and W leaves the elements of $\mathcal{H}$ fixed.

This theorem has numerous interesting consequences, of which a few will now be presented. Of course the trick is to find interesting positive-definite functions. In the case of the group of integers, there is a well-known connection between positive-definite functions and functions analytic in the disc with non-negative real part. With II this leads to the following theorem.

III. Theorem. Let D be the open unit disc in the complex plane, let $\mathcal{H}$ be a complex Hilbert space, and let $T: D \to \mathcal{B}(\mathcal{H})$ be analytic and satisfy $\operatorname{Re} T(z) \geq 0$ in D and $T(0) = I$. Then there is a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a unitary operator U on $\mathcal{K}$ such that

$$T(z) = P(I + zU)(I - zU)^{-1}$$

for all $z \in D$. Conversely, the function defined by this equation is analytic with $\operatorname{Re} T(z) \geq 0$ and $T(0) = I$.

Proof: Since T is analytic in D it admits a strongly convergent expansion $T(z) = \Sigma_{n=0}^{\infty} A_n z^n$ in D. Let A be the operator-valued function defined on the group of integers by $A(n) = \frac{1}{2} A_n$ for $n > 0$, $A(0) = I$, and $A(n) = \frac{1}{2} A^*_{-n}$ for $n < 0$. Then A is positive-definite, for if $x_k \in \mathcal{H}$ for $-n \leq k \leq n$, $y(\theta) = \Sigma_{k=-n}^{n} e^{-ik\theta} x_k$, and $0 < r < 1$, computation shows that

$$\frac{1}{2\pi}\int_0^{2\pi} (\operatorname{Re} T(re^{i\theta})y(\theta), y(\theta))d\theta = \sum_{k,\ell} r^{|k-\ell|}(A(k-\ell)x_k, x_\ell) .$$

Consequently there is a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a unitary operator U on $\mathcal{K}$ such that $A(n) = PU^n|\mathcal{H}$ for all n. Therefore

$$\begin{aligned} T(z) &= I + 2 \sum_{n=1}^{\infty} A(n)z^n \\ &= P(I + 2 \sum_{n=1}^{\infty} U^n z^n)\,|\,\mathcal{H} \\ &= P(I + zU)(I - zU)^{-1}\,|\,\mathcal{H} \ . \end{aligned}$$

For the converse, if $U = \int_0^{2\pi} e^{i\theta} dE_\theta$, then

$$(\operatorname{Re} T(z)x, x) = \int_0^{2\pi} \operatorname{Re}(1 + ze^{i\theta})(1 - ze^{i\theta})^{-1} d\|E_\theta x\|^2 \geq 0$$

since the integrand is non-negative for all $z \in D$ and all θ.

When $\mathcal{H}$ is one-dimensional, this result reduces to Herglotz's Theorem, which runs as follows: if f is analytic in D

and satisfies $\operatorname{Re} f \geq 0$ there and $f(0) \geq 0$, then

$$f(z) = \int_0^{2\pi} (e^{i\theta} + z)(e^{i\theta} - z)^{-1} d\mu(\theta)$$

for some positive finite Borel measure μ on $[0, 2\pi]$.

The existence of a strong unitary dilation for a contraction can be deduced from this theorem, in essentially the same way that Nagy [48] proved it. The key observation is that the conditions $\|T\| \leq 1$ and $\operatorname{Re}(I + zT)(I - zT)^{-1} \geq 0$ for $|z| < 1$ are equivalent. For if $x \in \mathcal{H}$ and $y = (I - zT)^{-1}x$, then

$$\begin{aligned}(\operatorname{Re}(I + zT)(I - zT)^{-1}x, x) &= \operatorname{Re}((I + zT)y, (I - zT)y) \\ &= \|y\|^2 - |z|^2 \|Ty\|^2,\end{aligned}$$

which implies the assertion. Therefore if T is a contraction, there is $\mathcal{K} \supset \mathcal{H}$ and a unitary operator U on $\mathcal{K}$ such that $(I + zT)(I - zT)^{-1} = P(I + zU)(I - zU)^{-1} | \mathcal{H}$. Expanding both sides and comparing coefficients gives $T^n = PU^n | \mathcal{H}$ for all $n \geq 0$.

A similar application was noticed recently by C. A. Berger [5]. For an operator T, let $w(T)$ denote the numerical radius $\sup \{|(Tx, x)| \mid \|x\| = 1\}$. Here the basic observation is that the conditions $w(T) \leq 1$ and $\operatorname{Re}(I - zT)^{-1} \geq 0$ for $|z| < 1$ are equivalent. To see this let $x \in \mathcal{H}$ and $y = (I - zT)^{-1}x$. Then $(\operatorname{Re}(I - zT)^{-1}x, x) = \operatorname{Re}(y, (I - zT)y) = \|y\|^2 - \operatorname{Re} z(Ty, y)$, which implies the assertion.

IV. Theorem (Berger). Let T be an operator on $\mathcal{H}$. Then $w(T) \leq 1$ if and only if there is a unitary operator U on a space $\mathcal{K} \supset \mathcal{H}$ such that $T^n = 2PU^n | \mathcal{H}$ for all $n > 0$.

Proof: By III and the above remarks the conditions

$$w(T) \leq 1$$

and

$$(I - zT)^{-1} = P(I + zU)(I - zU)^{-1} | \mathcal{H}$$

for $|z| < 1$ are equivalent. On expanding and equating coefficients the latter condition is seen to be equivalent to $T^n = 2PU^n | \mathcal{H}$ for $n > 0$.

One interesting consequence of IV is the following power inequality for the numerical radius.

COROLLARY. $w(T^m) \leq w(T)^m$ for any operator T and positive integer m.

Proof: By a change of scale it can be assumed that $w(T) \leq 1$, so that $T^n = 2PU^n | \mathcal{H}$ for all $n > 0$. The latter equation then holds with T replaced by T^m and U by U^m, so that $w(T^m) \leq 1$ by the converse part of IV.

This corollary has been generalized by Berger and Stampfli [6], as follows: if f is analytic in $|z| < 1$, continuous in $|z| \leq 1$, and $f(0) = 0$, and if $w(T) \leq 1$, then $w(f(T)) \leq \max |f(e^{i\theta})|$.

COROLLARY. If $w(T) \leq 1$, then $\|T^n\| \leq 2$ for all $n \geq 0$.

Theorem II can be refined considerably in the case of a locally compact abelian group, owing to the following description of the positive-definite functions on such a group.

V. **Bochner's Theorem.** If ϕ is a continuous positive-definite function on a locally compact abelian group G, there is a unique positive regular Borel measure μ on the dual group Γ such that for all $g \in G$,

$$\phi(g) = \int_\Gamma \gamma(g) d\mu(\gamma) \quad .$$

This was proved by Herglotz for the group of integers (see above), by Bochner for the real line, and by Weil in the general case. An easily accessible proof may be found in Rudin [42]. In view of the connection between positive-definite functions and unitary representations, it is not surprising that this theorem can be made to say something about unitary representations.

VI. **Stone's Theorem.** Let U be a weakly continuous unitary representation on a Hilbert space $\mathcal{H}$ of a locally compact abe-abelian group G. Then there is a unique regular spectral measure E on the Borel sets of the dual group Γ such that

$$(U(g)x, y) = \int_\Gamma \gamma(g) d(E(\gamma)x, y)$$

for all $x, y \in \mathcal{H}$ and $g \in G$.

The proof will only be sketched. (For background material on positive operator-valued measures, the reader may consult Berberian's excellent account [4], especially sections 1–6. A P.O. *measure* is a positive operator-valued function F on a σ-algebra of sets such that $(F(\cdot)x, x)$ is a measure for all x; a *spectral measure* is a projection-valued P.O. measure.) For each $x \in \mathcal{H}$ the function $\phi(g) = (U(g)x, x)$ is continuous and positive-definite, so there is a regular Borel measure μ_x on γ such that $(U(g)x, x) = \int_\Gamma \gamma(g)d\mu_x(\gamma)$. The polarization identity then provides a regular complex measure $\mu_{x,y}$ for each pair $x, y \in \mathcal{H}$ such that

$$(U(g)x, y) = \int_\Gamma \gamma(g)d\mu_{x,y}(\gamma) \quad .$$

This measure is unique, since any regular complex measure which annihilates the functions $\gamma \to \gamma(g)$ for all $g \in G$ is zero [42, p. 17]. The uniqueness now implies a standard way [21, p. 70] the existence of a unique regular spectral measure E such that $\mu_{x,y} = (E(\cdot)x, y)$ for all $x, y \in \mathcal{H}$. This concludes the sketch of the proof.

Before proceeding, it should be noticed that Stone's Theorem in the case of the group of integers amounts to the spectral theorem for unitary operators, and indicates how the latter theorem may be deduced from Herglotz's Theorem.

Now recall the statement of II, that a positive-definite operator-valued function A on a group G can be dilated to a unitary representation U: $A = PU \mid \mathcal{H}$, or equivalently, $(A(g)x, y) = (U(g)x, y)$ for all $x, y \in \mathcal{H}$. If G is locally compact and abelian, then by VI

$$(A(g)x, y) = \int_\Gamma \gamma(g)d(E(\gamma)x, y)$$

for a spectral measure E on Γ. Now the function defined by $F = PE \mid \mathcal{H}$ is a P.O. measure, and so the following has been proved.

VII. Theorem. If A is a positive-definite operator-valued function on a locally compact abelian group G, there is a unique regular P.O. measure F on the dual group Γ such that

$$(A(g)x, y) = \int_{\Gamma} \gamma(g) d(F(\gamma)x, y)$$

for all $x, y \in \mathcal{H}$ and $g \in G$.

Of course this result includes Bochner's Theorem. Although F is not a spectral measure, it has the advantage that an extension space is not used. The proof raises the following question: *can a* P.O. *measure always be dilated to a spectral measure?* In the context of locally compact abelian groups an easy affirmative answer is possible. For if F is a regular P.O. measure on Γ, the operator-valued function A on G defined by

$$(A(g)x, y) = \int_{\Gamma} \gamma(g) d(F(\gamma)x, y)$$

is easily seen to be positive-definite, so that as above,

$$(A(g)x, y) = \int_{\Gamma} \gamma(g) d(E(\gamma)x, y)$$

for a spectral measure E on an extension space. Comparing these equations gives $F = PE|\mathcal{H}$ as required. An "abstract" version of this result runs as follows:

VIII. Theorem. Let $\mathcal{H}$ be a Hilbert space, X a set, and $\mathcal{S}$ a ring of subsets of X with $\emptyset, X \in \mathcal{S}$, and let F be a function on $\mathcal{S}$ such that

(i) the values of F are positive operators on $\mathcal{H}$,

(ii) $F(\emptyset) = 0$ and $F(X) = I$, and

(iii) $F(S \cup T) = F(S) + F(T)$ if S and T are disjoint.

Then there is a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a function E on $\mathcal{S}$ such that the values of E are projections in $\mathcal{K}$, E satisfies (ii) and (iii), and $F = PE \mid \mathcal{H}$.

The argument is much like that used to prove II, and will only be outlined. Let $\mathcal{K}_1$ consist of all $\mathcal{S}$-measurable simple functions on X with values in $\mathcal{H}$. If $\xi, \eta \in \mathcal{K}_1$, and if $X = S_1 \cup \cdots \cup S_n$ is a measurable partition such that ξ and η are constant in each S_k, say with values x_k and y_k, then define

$$(\xi, \eta)_1 = \Sigma (F(S_k) x_k, y_k) .$$

As before $\mathcal{K}_0 = \{\xi \mid (\xi, \xi)_1 = 0\}$ is a subspace, $\mathcal{K}_1/\mathcal{K}_0$ is a pre-Hilbert space, and $\mathcal{K}$ is its completion. The constant functions provide an embedding of $\mathcal{H}$ into $\mathcal{K}$, and $E(S)\xi$ is defined to be equal to ξ on S and 0 on $X - S$.

Except as otherwise noted, the theorems of this section are due to Naimark [34, 35]. In [51] may be found a general dilation theorem of Nagy which includes both II and VIII, as well as an excellent discussion of many related topics.

SECTION 9

CONTRACTIVE SEMIGROUPS

Basic references for this subject are Hille-Phillips [29] and Dunford-Schwartz [16]. The appendix of Lax-Phillips [32] contains a brief and elegant account.

DEFINITION. A family $T = \{T(t) | t \geq 0\}$ of operators on a Hilbert space $\mathcal{H}$ is called a *semigroup* if

$$T(0) = I$$

and

$$T(s+t) = T(s)T(t) \text{ for all } s, t \geq 0 .$$

It is *uniformly continuous* if $t \to T(t)$ is continuous in the operator norm, and *strongly continuous* if $t \to T(t)x$ is continuous in the norm of $\mathcal{H}$ for all $x \in \mathcal{H}$. (Because of the semigroup property it is sufficient to assume continuity at $t = 0$.) If $T(t)$ is a contraction for all $t \geq 0$, the semigroup is a *contractive semigroup*.

As examples we mention the following:

1. Let A be an operator and $T(t) = e^{tA}$. It is shown below that these are precisely the uniformly continuous semigroups.

2. Unitary semigroups. A semigroup such that all $T(t)$ are unitary amounts to a unitary representation of the group

of real numbers, and these are described by Stone's Theorem §8.VI.

3. Translation semigroups. Let $\mathcal{K}$ be a separable Hilbert space, and denote by $L^2(R^+,\mathcal{K})$ the Hilbert space of weakly measurable functions f from $[0,\infty)$ into $\mathcal{K}$ such that

$$\|f\|^2 = \int_0^\infty \|f(s)\|^2 ds < \infty .$$

The (backward) *translation semigroup* is defined on $L^2(R^+,\mathcal{K})$ by $(B(t)f)(s) = f(s+t)$. It is easily seen that

$$(B(t)^*f)(s) = \begin{cases} 0 \text{ for } 0 \leq s < t \\ f(s-t) \text{ for } t \leq s . \end{cases}$$

This semigroup, consisting of isometries, is the forward translation semigroup. The bilateral translation semigroup, defined analogously on $L^2(R,\mathcal{K})$, is unitary. These semigroups are strongly (but not uniformly) continuous.

4. The restriction to an invariant subspace of a semigroup is again a semigroup, as is a direct sum of semigroups. The adjoint of a semigroup is a semigroup; it is easy to see that the adjoint of a strongly continuous contractive semigroup is strongly continuous.

I. Theorem. $\{T(t)\}$ is a uniformly continuous semigroup on $\mathcal{H}$ if and only if there is $A \in \mathcal{B}(\mathcal{H})$ with $T(t) = e^{tA}$ for all $t \geq 0$.

Proof: For $B \in \mathcal{B}(\mathcal{H})$, e^B is defined by means of the usual uniformly convergent power series. If B and C commute, the estimate

$$\|e^B - e^C\| \leq e^M \|B - C\|$$

follows readily, where $M = \max\{\|B\|, \|C\|\}$. Hence e^{tB} is uniformly continuous. That it is a semigroup is proved just as in the scalar case.

If $\{T(t)\}$ is a uniformly continuous semigroup, then

$$\| \frac{1}{t} \int_0^t T(s)ds - I\| \to 0 \text{ as } t \to 0+$$

because the integrand is continuous, and so there is $a > 0$ such that $\int_0^t T(s)ds$ is invertible for $0 < t \leq a$. If $0 < h \leq t \leq a$,

$$\begin{aligned} \frac{1}{h}(T(h)-I) \int_0^t T(s)ds &= \frac{1}{h}\left\{ \int_0^t T(h)T(s)ds - \int_0^t T(s)ds \right\} \\ &= \frac{1}{h}\left\{ \int_h^{t+h} T(s)ds - \int_0^t T(s)ds \right\} \\ &= \frac{1}{h}\left\{ \int_t^{t+h} T(s)ds - \int_0^h T(s)ds \right\} \end{aligned}$$

which converges in the operator norm as $h \to 0+$ (to $T(t)-I$). Hence $A = \lim 1/h(T(h)-I)$ exists in $\mathcal{B}(\mathcal{H})$, and therefore

$$\frac{d}{dt}T(t) = \lim \frac{1}{h}(T(t+h) - T(t)) = AT(t) \quad .$$

The semigroup $T_1(t) = e^{tA}$ satisfies the same differential equation, which implies that $T(s-t)T_1(t)$ is constant on $[0, s]$ for any $s > 0$, and consequently that $T(s) = T_1(s)$.

It should be noticed that the theorem makes sense and the proof is valid with $\mathcal{B}(\mathcal{H})$ replaced by any Banach algebra with identity.

The remainder of this section will be devoted to elucidating the structure of strongly continuous contractive semigroups. This may be thought of as the continuous analogue of the study carried out in sections 1 and 6 for (the powers of) a single contration. Two methods will be employed: adapt the methods of the discrete case, and extend the procedures used in the proof of I above. In each the notion of the infinitesimal generator plays an important role.

DEFINITION. The *infinitesimal generator* A of the semigroup $\{T(t)\}$ is defined by

$$Ax = \lim_{h \to 0+} \frac{1}{h} (T(h) - I)x$$

on the set $\mathfrak{D} = \mathfrak{D}(A)$ of vectors x for which this limit exists.

It is clear that $\mathfrak{D}$ is a linear manifold and that A is linear on $\mathfrak{D}$. In general $\mathfrak{D}$ will be a proper submanifold and A will fail to be bounded. For the backward translation semigroup of example 3, $\mathfrak{D}$ consists of the differentiable functions $f \in L^2$ for which $f' \in L^2$, and A is differentiation. The following basic properties of A will be needed.

II. Lemma. Let $\{T(t)\}$ be a strongly continuous semigroup with infinitesimal generator. A. Then:

1. A commutes with $T(t)$ for all $t \geq 0$;
2. $\frac{d}{dt} T(t)x = T(t)Ax$ and $\frac{d}{dt} \| T(t)x\|^2 = 2\text{Re}\,(AT(t)x, T(t)x)$ for all $x \in \mathfrak{D}(A)$;
3. $\int_0^a T(s)x\,ds \in \mathfrak{D}(A)$ for all $x \in \mathcal{H}$ and $a \geq 0$, and $A \int_0^a T(s)x\,ds = (T(a) - I)x$; and
4. A is closed and densely defined.

Proof: The first statement is that if $x \in \mathfrak{D}(A)$ then $T(t)x \in \mathfrak{D}(A)$ and $AT(t)x = T(t)Ax$, and this is clear from the definition. For the second, if $x \in \mathfrak{D}(A)$ then

$$\frac{1}{h}(T(t+h)-T(t))x = T(t)\frac{1}{h}(T(h)-I)x \to T(t)Ax$$

$$\frac{1}{-h}(T(t-h)-T(t))x = T(t-h)\frac{1}{h}(T(h)-I)x \to T(t)Ax$$

as $h \to 0+$. The second part of 2 follows from the first and the fact that

$$(f(t),g(t))' = (f'(t),g(t)) + (f(t),g'(t)) \quad .$$

Just as in the proof of I,

$$\frac{1}{h}(T(h)-I)\int_0^a T(s)x\,ds \to (T(a)-I)x\ ,$$

which is 3. Since $1/a \int_0^a T(s)x\,ds \to x$ as $a \to 0+$, 3 implies that $\mathfrak{D}(A)$ is dense. To see that A is closed, notice first that integrating the relation in 2 gives

$$\int_0^t T(s)Ax\,ds = (T(t)-I)x$$

for $x \in \mathfrak{D}(A)$. If $x_n \in \mathfrak{D}(A)$, $x_n \to x$, and $Ax_n \to y$, letting $n \to \infty$ in $\int_0^t T(s)Ax_n ds = (T(t)-I)x_n$ gives

$$\int_0^t T(s)y\,ds = (T(t)-I)x \quad .$$

Dividing by t and letting $t \to 0+$ gives $Ax = y$ as required.

III. Theorem. Any strongly continuous contractive semigroup can be extended to a strongly continuous coisometric semigroup.

Proof: Let $\{T(t)\}$ be a strongly continuous contractive semigroup and A its infinitesimal generator. The hermitian symmetric bilinear form defined on $\mathfrak{D} = \mathfrak{D}(A)$ by

$$(x, y)_1 = -(Ax, y) - (x, Ay)$$

is positive-definite by II(2) since $\|T(t)x\|$ is non-increasing. It follows that $\mathfrak{N} = \{x \in \mathfrak{D} \mid (x, x)_1 = 0\}$ is a submanifold of $\mathfrak{D}$, and that the bilinear form induced on $\mathfrak{D}/\mathfrak{N}$ makes it a pre-Hilbert space. Let $\mathcal{K}$ be the completion. The positive contractions $T(t)^*T(t)$ are decreasing and therefore converge strongly to a positive contraction. Let C be the positive square root of this contraction, so that $\lim \|T(t)x\|^2 = \|Cx\|^2$, and let $\mathfrak{L} = (C\mathcal{H})^-$. Now define

$$\Sigma\colon\ \mathfrak{D} \to L^2(R^+, \mathcal{K}) \oplus \mathfrak{L}$$

by $\Sigma x = Wx \oplus Cx$, where $(Wx)(t) = T(t)x$. Then

$$\begin{aligned} \|Wx\|^2 &= \lim_{n\to\infty} \int_0^n \|T(t)x\|_1^2\, dt \\ &= -\lim \int_0^n 2\,\mathrm{Re}(AT(t)x, T(t)x)dt \\ &= -\lim \int_0^n \frac{d}{dt}\|T(t)x\|^2\, dt \\ &= \|x\|^2 - \lim \|T(n)x\|^2 \\ &= \|x\|^2 - \|Cx\|^2 \end{aligned}$$

by II(2), so that Σ is an isometry on $\mathfrak{D}$. Since $\mathfrak{D}$ is dense in $\mathcal{H}$, Σ has a unique isometric extension to all of $\mathcal{H}$. Now if $x \in \mathfrak{D}$ then $T(s)x \in \mathfrak{D}$ and $\Sigma T(s)x = WT(s)x \oplus CT(s)x$. But

$$(WT(s)x)(t) = T(t)T(s)x = T(t+s)x = (B(s)Wx)(t) ,$$

where B is the translation semigroup of example 3. In addition, $V(s)$: $Cx \to CT(s)x$ is a well-defined isometry on $C\mathfrak{D}$, since

$$\|CT(s)x\|^2 = \lim \|T(t)T(s)x\|^2 = \lim \|T(t)x\|^2 = \|Cx\|^2 ,$$

and so $V(s)$ has a unique isometric extension to all of $\mathfrak{L}$. Therefore

$$\Sigma T(s)x = B(s)Wx \oplus V(s)Cx = (B(s) \oplus V(s))\Sigma x$$

for all $x \in \mathfrak{D}$, so that by continuity

$$\Sigma T(s) = (B(s) \oplus V(s))\Sigma$$

on $\mathcal{H}$. This implies that $\mathfrak{R} = \Sigma\mathcal{H}$ is an invariant subspace for the semigroup $B \oplus V$, and that T and $(B \oplus V) \mid \mathfrak{R}$ are unirarily equivalent. (Since $V(s)C = CT(s)$ it follows easily that V is a strongly continuous semigroup.)

The final step in the proof is, as in the discrete case, the extension of the isometric semigroup V to a unitary semigroup. This is accomplished by means of the following continuous analogue of the Wold decomposition §1.II, due to J. L. B. Cooper [14]. The ingenious proof given below was discovered by James Deddens. Other interesting proofs may be found in Masani [33] and Sz.-Nagy [50].

IV. **Theorem.** Let $V = \{V(t)\}$ be a strongly continuous isometric semigroup on $\mathcal{H}$. Then there are Hilbert spaces $\mathcal{K}$ and

$\mathcal{L}$ and a strongly continuous unitary semigroup $U = \{U(t)\}$ on $\mathcal{L}$, such that V is unitarily equivalent to $B^* \oplus U$, where B^* is the forward translation semigroup on $L^2(R^+, \mathcal{K})$ (c.f. example 3).

Proof: Apply the above reasoning to the strongly continuous coisometric semigroup $T(t) = V(t)^*$. Thus there are Hilbert spaces $\mathcal{K}$ and $\mathcal{L}$, a strongly continuous isometric semigroup V_1 on $\mathcal{L}$, and an isometry Σ from $\mathcal{H}$ into $L^2(R^+, \mathcal{K}) \oplus \mathcal{L}$ such that

$$\Sigma T(s) = (B(s) \oplus V_1(s))\Sigma \quad .$$

In this case it will be shown that V_1 is unitary and that $\Sigma \mathcal{H} = L^2(R^+, \mathcal{K}) \oplus \mathcal{L}$; the theorem will then follow on taking adjoints.

Since T is coisometric the operators $T(t)^* T(t)$ are projections, and therefore so is C. The equation $T(t)^* C^2 T(t) = C^2$ implies that C^2 commutes with $T(t)$, again because $T(t)$ is a coisometry, and hence that C commutes with $T(t)$. Thus $\mathcal{L} = C\mathcal{H}$ reduces T and $V_1 = T \mid \mathcal{L}$; since V_1 is isometric and $T \mid \mathcal{L}$ is coisometric, both must be unitary.

Let P be the projection of $L^2(R^+, \mathcal{K}) \oplus \mathcal{L}$ on $\mathcal{L}$. Then $P\Sigma = C$, so that for $x \in \mathcal{L}$, $\Sigma x = f \oplus Cx = f \oplus x$ for some $f \in L^2(R^+, \mathcal{K})$. But $f = 0$ since Σ is an isometry, and this gives $\Sigma \mathcal{H} = \mathcal{M} \oplus \mathcal{L}$ where $\mathcal{M}$ is a subspace of $L^2(R^+, \mathcal{K})$ invariant for B. Now $B \mid \mathcal{M}$ and $T \mid (I - C)\mathcal{H}$ are unitarily equivalent. Since the restriction of a coisometry to an invariant subspace is a coisometry if and only if the subspace reduces, it follows that $\mathcal{M}$ reduces B. To get $\mathcal{M} = L^2(R^+, \mathcal{K})$ it will suffice to show that $\mathcal{M}$ contains the step functions with range in $\mathfrak{D}$. Since $\mathfrak{M}$ reduces we need only consider functions of the form $f = \phi_{[0,a]}x$, where $\phi_{[0,a]}$ is the characteristic function of $[0,a]$, $a > 0$, and $x \in \mathfrak{D}$. Now

$$g = WT(b)x - B(b)^* WT(2b)x \in \mathfrak{M} \text{ for } b \geq 0$$

(where W is as on p. 82), and

$$g(t) = \begin{cases} T(t+b)x, & t \leq b \\ 0, & b < t \end{cases} .$$

Let $\varepsilon > 0$. Since

$$\begin{aligned} \|(I - T(s))x\|_1^2 &= -2\,\mathrm{Re}\,(A(I - T(s))x, (I - T(s))x) \\ &= -2\,\mathrm{Re}\,((I - T(s))Ax, (I - T(s))x) \end{aligned}$$

converges to 0 as $s \to 0+$, there is an integer $N > 0$ such that $\|x - T(s)x\|_1^2 < \varepsilon/a$ for $s < 2a/N$. If $b = a/N$ and $h = \Sigma_{n=1}^{N} B((n-1)b)^* g$, then $h \in \mathfrak{M}$ and

$$\begin{aligned} \|h - f\|^2 &= \int_0^a \|h(t) - x\|_1^2 \, dt \\ &= \sum_{n=1}^{N} \int_{(n-1)b}^{nb} \|g(t - (n-1)b) - x\|_1^2 \, dt \\ &= N \int_0^b \|T(t+b)x - x\|_1^2 \, dt \\ &< Nb(\varepsilon/a) = \varepsilon \quad . \end{aligned}$$

Hence $f \in \mathfrak{M}$ and the proof is complete.

COROLLARY. Any strongly continuous isometric semigroup can be extended to a strongly continuous unitary semigroup.

COROLLARY. Any strongly continuous contractive semigroup can be dilated to a strongly continuous unitary semigroup.

REMARKS. 1. The effect of these results is to reduce the study of contractive semigroups to the study of invariant subspaces of coisometric semigroups. In the case of the translation semigroup B, a further reduction is possible, just as in the discrete case: the Fourier transforms of the invariant subspaces of B may be expressed in terms of certain operator-valued functions which are analytic in a half-plane [30, 31].

2. The uniqueness situation is the same as in the discrete case §6.I. It is also worth noting that the invariant subspace $\Sigma\mathcal{H}$ in the proof of III is full. For the preceding step-function argument actually shows that $(P\Sigma\mathcal{H})^-$ is always a full invariant subspace of $L^2(R^+, \mathcal{K})$. In addition, if U on $\mathcal{F}$ is the unitary extension of V obtained from Cooper's Theorem, and if Q is the projection on $\mathcal{F}$, then it is almost obvious that $(Q\Sigma\mathcal{H})^-$ is a full invariant subspace of $\mathcal{F}$. These facts imply that $\Sigma\mathcal{H}$ is full.

Exercises. 1. The translation semigroups are strongly continuous.

2. The adjoint of a strongly continuous contractive semigroup is strongly continuous.

3. If f is a continuous mapping from $[0, \infty)$ into a Banach space, then

$$\frac{1}{t}\int_a^{a+t} f(s)ds \to f(a) \text{ as } t \to 0.$$

4. If C is a coisometry and $\mathfrak{M}$ an invariant subspace, then $C \mid \mathfrak{M}$ is a coisometry if and only if $\mathfrak{M}$ reduces C.

5. A strongly continuous contractive semigroup is unitarlly equivalent to a part of a backward translation semigroup if and only if $T(t) \to 0$ strongly as $t \to \infty$.

6. A semigroup T is unitarily equivalent to a backward translation semigroup if and only if it is strongly continuous, coisometric, and $T(t) \to 0$ strongly as $t \to \infty$.

In the remainder of this section the second method mentioned above will be considered. Here the procedure is to characterize the infinitesimal generators, and to develop ways of recovering the semigroup from its generator. Again only strongly continuous contractive semigroups will be considered.

DEFINITION. A linear transformation A in a Hilbert space $\mathfrak{H}$ is called *accretive*[1] if it is densely defined and if

$$\operatorname{Re}(Ax, x) \leq 0 \quad \text{for all} \quad x \in \mathfrak{D}(A),$$

and *maximal accretive* if it is accretive and admits no proper accretive extension (in $\mathfrak{H}$).

If A is accretive, then for any $x \in \mathfrak{D}(A)$

$$\|(A \pm I)x\|^2 = \|Ax\|^2 + \|x\|^2 \pm \operatorname{Re}(Ax, x)$$

and consequently

$$\|(A + I)x\|^2 \leq \|Ax\|^2 + \|x\|^2 \leq \|(A - I)x\|^2 ,$$

so that $A - I$ is one-to-one and $S = (A + I)(A - I)^{-1}$

---

* See the remark on terminology on p. 93.

is a contraction defined on $(A-I)\,\mathfrak{D}(A)$. This contraction will be referred to as the *Cayley transform* of A. Since $S-I = 2(A-I)^{-1}$ and $S+I = 2A(A-I)^{-1}$, it follows that $S-I$ is one-to-one and

$$(S+I)(S-I)^{-1} = A\ .$$

This implies that if A_1 and A_2 are accretive with Cayley transforms S_1 and S_2, then A_1 is a (proper) extension of A_2 if and only if S_1 is a (proper) extension of S_2.

V. Let A be a densely defined linear transformation in $\mathcal{H}$. The following conditions are equivalent:

1. A is maximal accretive;
2. A is accretive and $(A-I)\,\mathfrak{D}(A) = \mathcal{H}$;
3. $A = (S+I)(S-I)^{-1}$ for some everywhere defined contraction S of which 1 is not an eigenvalue; and
4. A and A^* are closed and accretive.

Proof: That 1 and 2 are equivalent and imply 3 is clear from the preceding discussion. If $A = (S+I)(S-I)^{-1}$ is as in 3, and $x = (S-I)y \in \mathfrak{D}(A)$, then

$$\operatorname{Re}(Ax, x) = \operatorname{Re}((S+I)y, (S-I)y) = \|Sy\|^2 - \|y\|^2 \leq 0$$

so that A is accretive; it is maximal because S is everywhere defined. Thus 3 implies 1.

Again consider $A = (S+I)(S-I)^{-1}$ as in 3. If $S^*x = x$, then

$$\begin{aligned}\|Sx - x\|^2 &= \|Sx\|^2 - 2\operatorname{Re}(Sx, x) + \|x\|^2\\ &= \|Sx\|^2 - 2\operatorname{Re}(x, S^*x) + \|x\|^2\\ &= \|Sx\|^2 - \|x\|^2 \leq 0\ ,\end{aligned}$$

so that $Sx = x$ and $x = 0$. Hence 1 is not an eigenvalue of S^*, and therefore

$$\begin{aligned}(S^* + I)(S^* - I)^{-1} &= -I + 2(S^* - I)^{-1} \\ &= [-I + 2(S-I)^{-1}]^* \\ &= A^*\end{aligned}$$

is maximal accretive. To prove 4 it will now suffice to show that a maximal accretive transformation is closed. Since $(A-I)^{-1}$ is bounded with domain $(A-I)\mathfrak{D}(A)$, an accretive transformation A is closed if and only if $(A-I)\mathfrak{D}(A)$ is closed. Hence maximal accretive transformations are closed by 2.

Finally assume 4. If y is orthogonal to $(A-I)\mathfrak{D}(A)$, then $A^*y = y$, so that $y = 0$ since A^* is accretive. Therefore $(A-I)\mathfrak{D}(A)$ is dense in $\mathcal{H}$. Since A is closed this gives $(A-I)\mathfrak{D}(A) = \mathcal{H}$ by the above remark.

VI. An accretive linear transform has a maximal accretive extension.

Proof: Let A be accretive with Cayley transform S. By the foregoing it is sufficient to extend S to an everywhere defined contraction S_0 of which 1 is not an eigenvalue. Define S_0 on the closure of $(A-I)\mathfrak{D}(A)$ by continuity and on the orthogonal complement by $S_0 = 0$. If $S_0 x = x$, then as in the preceding proof $S_0^* x = x$, so that for any $y \in \mathfrak{D}(A)$,

$$\begin{aligned}(x, (A-I)y) &= (S_0^* x, (A-I)y) \\ &= (x, S(A-I)y) \\ &= (x, (A+I)y) \quad ,\end{aligned}$$

$(x, y) = 0$, and $x = 0$ because $\mathfrak{D}(A)$ is dense.

VII. The infinitesimal generator of any strongly continuous contractive semigroup is maximal accretive.

Proof: Since $\|T(t)x\|$ is nonincreasing, A is accretive by II(2). The generator of the semigroup $\{e^{-t}T(t)\}$ is $A-I$, so that by II(3)

$$e^{-t}T(t)x - x = (A-I)\int_0^t e^{-s}T(s)x\,ds$$

for any $x \in \mathcal{H}$. Since $y = \int_0^\infty e^{-s}T(s)x\,ds$ converges and $A-I$ is closed, this gives $(A-I)y = -x$. Thus $(A-I)\mathfrak{D}(A) = \mathcal{H}$, so that A is maximal by V.

REMARK. $(I-A)^{-1}$ is given by the Laplace transform $\int_0^\infty e^{-s}T(s)ds$. More generally,

$$(\lambda I - A)^{-1} = \int_0^\infty e^{-\lambda s}T(s)ds \text{ for } \operatorname{Re}\lambda > 0 \ .$$

The next step is the converse: any maximal accretive transformation A is the generator of a unique strongly continuous contractive semigroup. The semigroup in question is simply e^{tA}; the difficulty is to make sense of the exponential. There are several ways to do this, in each of which the idea is to construct approximating semigroups e^{tB}; where B is a *bounded* approximation to A. The following is a sketch of the method of Hille and Yosida.

For $\lambda > 0$ it follows just as above that $\lambda I - A$ has an everywhere defined inverse R_λ which is bounded by $1/\lambda$. The operators

$$B_\lambda = \lambda^2 R_\lambda - \lambda I$$

converge strongly to A on $\mathfrak{D}(A)$ as $\lambda \to \infty$, and the semigroups $T_\lambda(t) = \exp(tB_\lambda)$ are strongly continuous and consist of contractions. Since

$$T_\lambda(t) - T_\mu(t) = \int_0^1 \frac{d}{ds} T_\lambda(ts)T_\mu(t(1-s))ds$$
$$= \int_0^1 tT_\lambda(ts)T_\mu(t(1-s))(B_\lambda - B_\mu)ds \ ,$$

it follows that

$$\| T_\lambda(t)x - T_\mu(t)x\| \leq t\|B_\lambda x - B_\mu x\|$$

and hence that a limiting strongly continuous contractive semigroup $T(t)$ exists. If A_1 is its generator, then for $x \in \mathfrak{D}(A)$

$$(T_\lambda(h) - I)x = B_\lambda \int_0^h T_\lambda(s)x\,ds = \int_0^h T_\lambda(s)B_\lambda x\,ds \ ;$$

letting $\lambda \to \infty$ this gives

$$(T(h) - I)x = \int_0^h T(s)Ax\,ds$$

which implies that $A \subset A_1$. Equality follows since A is maximal accretive and A_1 is accretive (V). Uniqueness may be proved just as in I. This proves

VIII. Theorem. A maximal accretive operator is the infinitesimal generator of a unique strongly continuous contractive semigroup.

IX. If A is the infinitesimal generator of the strongly continuous contractive semigroup $\{T(t)\}$, then A^* is the infinitesimal generator of $\{T(t)^*\}$.

Proof: Let B be the generator of $\{T(t)^*\}$. If $x \in \mathfrak{D}(A)$ and $y \in \mathfrak{D}(B)$, then

$$(Ax, y) = \lim \frac{1}{h}((T(h)-I)x, y)$$

$$= \lim \frac{1}{h}(x, (T(h)^* - I)y) = (x, By) \quad ,$$

so that $B \subset A^*$. Equality follows since B is maximal accretive and A^* is accretive (by V and VII).

To properties of a semigroup T there are corresponding properties of the infinitesimal generator A and of the Cayley transform $S = (A+I)(A-I)^{-1}$ (S is called the *cogenerator* of T). For example:

X. Let T be a strongly continuous contractive semigroup with infinitesimal generator A and cogenerator S. The following statements are equivalent:

1. T is isometric;
2. $A \subset -A^*$ (equivalently, iA is symmetric);
3. $\mathrm{Re}(Ax, x) = 0$ for all $x \in \mathfrak{D}(A)$; and
4. S is an isometry.

Proof: If T is isometric, then for $x \in \mathfrak{D}(A)$

$$(T(h)^* - I)x = -T(h)^*(T(h) - I)x$$

so that $A \subset -A^*$ (using IX). Then

$$\mathrm{Re}(Ax, x) = -\mathrm{Re}(A^*x, x) = -\mathrm{Re}(Ax, x)$$

and consequently $\operatorname{Re}(Ax, x) = 0$. Since

$$\|(A \pm I)x\|^2 = \|Ax\|^2 + \|x\|^2 \pm 2\operatorname{Re}(Ax, x) \quad ,$$

S is an isometry if and only if $\operatorname{Re}(Ax, x) = 0$ for all $x \in \mathfrak{D}(A)$. Finally, if the latter holds, then $d/dt \|T(t)x\|^2 = 0$ for all $x \in \mathcal{H}$ by II(2), and so T is isometric.

COROLLARY. The following are equivalent: T is unitary, A is skew-adjoint, and S is unitary.

These facts may be used to obtain an alternative proof of the extension theorem III. For if S_1 is the minimal coisometric extension of the contraction S (§6.I), it is easy to see that 1 is not an eigenvalue of S_1, so that S_1 is the cogenerator of a coisometric semigroup T_1 which extends T. A proof of Cooper's Theorem along these lines can also be given [54].

REMARK. Let T be a strongly continuous contractive semigroup with generator A and cogenerator S. Let $R = (I - S^*S)^{1/2}$ and $\mathfrak{R} = (R\mathcal{H})^-$. Then the coefficient space $\mathcal{K}$ employed in the proof of the extension theorem III is isomorphic to $\mathfrak{R}$. For if $x \in \mathfrak{D}(A)$, one computes easily that

$$\|R(A - I)x\|^2 = 2\|x\|_1^2 \quad ,$$

so that $x \to (1/\sqrt{2})R(A - I)x$ induces the desired isometry.

REMARK ON TERMONOLOGY. Many authors prefer to define the infinitesimal generator by the equation $Bx = \lim_{h \to 0+} (i/h)(I - T(h))x$. Accretive then means $\operatorname{Im} Bx, x) \geq 0$, and the Cayley transform is defined as $S = (iI - B)(iI + B)^{-1}$ on $(iI + B)\mathfrak{D}(B)$.

Exercises. 1. Let T be a strongly continuous contractive semigroup with cogenerator S. A subspace $\mathfrak{M}$ is invariand for T if and only if it is invariant for S, and in this case $S \mid \mathfrak{M}$ is the cogenerator of $T \mid \mathfrak{M}$.

2. An everywhere defined accretive linear transformation is bounded.

3. Complete the proof of VIII.

SECTION 10

HYPERINVARIANT SUBSPACES

A subspace $\mathfrak{M}$ of a Hilbert space $\mathcal{H}$ is said to be *hyperinvariant* for an operator T on $\mathcal{H}$ if $S\mathfrak{M} \subset \mathfrak{M}$ for all operators S that commute with T. This concept is related to invariant subspaces for algebras of operators: the hyperinvariant subspaces of T are the invariant subspaces of the (strongly closed) operator algebra

$$\{T\}' = \{S \mid ST = TS\} .$$

Of course a scalar operator $T = cI$ has only the trivial hyperinvariant subspaces $\{0\}$ and $\mathcal{H}$. When $\mathcal{H}$ is finite-dimensional the converse assertion is true: a nonscalar operator has nontrivial hyperinvariant subspaces. More generally, a well-known result of Burnside asserts that the only algebra of $n \times n$ matrices without nontrivial invariant subspaces is the algebra of all $n \times n$ matrices. In general the question is open, even for compact operators.

Problem. Does every nonscalar (compact) operator have a nontrivial hyperinvariant subspace?

The corresponding problem for operator algebras is:

Problem. Is $\mathfrak{B}(\mathcal{H})$ the only strongly closed operator algebra without nontrivial invariant subspaces?

An affirmative answer for the latter problem (which seems quite unlikely) would imply an affirmative answer for the former, and hence for the invariant subspace problem. This section is concerned with several existence theorems for hyperinvariant subspaces. Beyond this, very little is known concerning these interesting and important problems.

Before proceeding, consider the case of a normal operator N. Because the spectral projections for N commute with any operator in $\{N\}'$, the spectral subspaces are hyperinvariant. On the other hand, since $A \in \{N\}'$ implies $A^* \in \{N\}'$ (Fuglede's Theorem), the projection P on a hyperinvariant subspace commutes with any operator in $\{N\}'$. It is well-known that when $\mathcal{H}$ is separable, such projections must be spectral. Thus in general it may be useful to think of the hyperinvariant subspaces as assuming the role of the spectral subspaces.

The following result, due to Rosenthal and Stampfli, shows that certain invariant subspaces must be hyperinvariant solely by virtue of their position in the lattice of all invariant subspaces.

I. Theorem. Let $\mathcal{S}$ be a countable family of invariant subspaces for an operator T, with the property that for any invariant subspaces $\mathfrak{M} \in \mathcal{S}$ and $\mathfrak{N} \notin \mathcal{S}$, either $\mathfrak{M} \subset \mathfrak{N}$ or $\mathfrak{N} \subset \mathfrak{M}$. Then $\mathcal{S}$ consists of hyperinvariant subspaces.

Proof: Observe first that for any operator S and $|\lambda| > \|S\|$, the operators S and $(S-\lambda I)^{-1}$ have the same invariant subspaces. For the series $-\Sigma\lambda^{-n-1}S^n$ converges to $(S-\lambda I)^{-1}$ in the operator norm, and so $S\mathfrak{M} \subset \mathfrak{M}$ implies

$$(S-\lambda I)^{-1}\mathfrak{M} \subset \mathfrak{M} .$$

On the other hand, if $(S-\lambda I)^{-1}\mathfrak{M} \subset \mathfrak{M}$ then $\mathfrak{M} \subset (S-\lambda I)\mathfrak{M}$. If the inclusion were proper there would be a unit vector $x \in \mathfrak{M}$ orthogonal to $(S-\lambda I)x$, so that

$$0 = ((S-\lambda I)x, x) \geq |\lambda| - |(Sx, x)| \geq |\lambda| - \|S\| > 0 \ ,$$

a contradiction. Thus $\mathfrak{M} = (S-\lambda I)\mathfrak{M}$ and $S\mathfrak{M} \subset \mathfrak{M}$.

Now suppose that S commutes with T and let $\mathfrak{M} \in \mathcal{S}$. Then the subspaces $(S-\lambda I)\mathfrak{M}$, $|\lambda| > \|S\|$, are invariant for T. If $(S-\lambda I)\mathfrak{M} \notin \mathcal{S}$ for some $|\lambda| > \|S\|$, then by hypothesis $\mathfrak{M} \subset (S-\lambda I)\mathfrak{M}$ or $(S-\lambda I)\mathfrak{M} \subset \mathfrak{M}$, so that $(S-\lambda I)^{-1}\mathfrak{M} \subset \mathfrak{M}$ or $(S-\lambda I)\mathfrak{M} \subset \mathfrak{M}$, and in either case $S\mathfrak{M} \subset \mathfrak{M}$. If $(S-\lambda I)\mathfrak{M} \in \mathcal{S}$ for all $|\lambda| > \|S\|$, then $(S-\lambda_1 I)\mathfrak{M} = (S-\lambda_2 I)\mathfrak{M}$ for some $\lambda_1 \neq \lambda_2$ since $\mathcal{S}$ is countable, and therefore

$$\mathfrak{M} = (S-\lambda_2 I)^{-1}(S-\lambda_1 I)\mathfrak{M} = (I+(\lambda_2-\lambda_1)(S-\lambda_2 I)^{-1})\mathfrak{M} \ ,$$

so that again $S\mathfrak{M} \subset \mathfrak{M}$. Hence $\mathfrak{M}$ is hyperinvariant.

COROLLARY. If the invariant subspaces of T are countable in number, then every invariant subspace is hyperinvariant.

COROLLARY. If $\mathfrak{M}$ is an invariant subspace of T that is comparable with every other invariant subspace of T, then $\mathfrak{M}$ is hyperinvariant.

COROLLARY. If the invariant subspaces of T form a chain, then every invariant subspace is hyperinvariant.

An operator such that the invariant subspaces form a chain (i.e., for any invariant subspaces $\mathfrak{M}$ and $\mathfrak{M}$, either $\mathfrak{M} \subset \mathfrak{N}$ or $\mathfrak{N} \subset \mathfrak{M}$) is called *unicellular*. When $\mathcal{H}$ is finite-dimensional, the unicellular operators are those of the form $\lambda I + N$, where N is cyclic and nilpotent (in other words, cyclic with one point spectrum). For arbitrary $\mathcal{H}$, it is not known whether the spectrum of a unicellular operator must reduce to a point. However [40]:

II. A unicellular operator on a separable Hilbert space has a cyclic vector.

Proof: Let T be unicellular and let $\{\mathfrak{M}_a \mid a \in A\}$ be the family of all proper invariant subspaces of T. The problem is to show that $\cup \mathfrak{M}_a \neq \mathcal{H}$. Let $\{x_i\}$ be a countable dense subset of $\cup \mathfrak{M}_a$, and let $x_i \in \mathfrak{M}_{a(i)}$. If $\mathfrak{M}_\beta \subset \cup \mathfrak{M}_{a(i)}$ is false for some β, then $\cup \mathfrak{M}_{a(i)} \subset \mathfrak{M}_\beta$ since $\{\mathfrak{M}_a\}$ is a chain, and then $\cup \mathfrak{M}_a \subset \mathfrak{M}_\beta \neq \mathcal{H}$. In the other case $\cup \mathfrak{M}_a = \cup \mathfrak{M}_{a(i)} \neq \mathcal{H}$ by the Baire Category Theorem.

We now discuss several examples of unicellular operators. The chain of invariant subspaces of the first will be shown to have the order type of a closed interval, and that of the second the order type of the positive integers with $+\infty$ adjoined. It is not known whether the order type of the integers with $\pm\infty$ adjoined occurs for some unicellular operator.

The *Volterra operator* is defined on $L^2[0,1]$ with Lebesgue measure by $(Vf)(t) = \int_0^t f$. It is clear that the subspaces $L^2[a,1]$, $0 \leq a \leq 1$, are invariant, and it will be shown that these are the only invariant subspaces. The proof we give is due to G. K. Kalisch, as simplified by M. Schreiber [45]. The formula

$$(V^n f)(t) = \frac{1}{(n-1)!} \int_0^t (t-s)^{n-1} f(s)\,ds$$

will be needed. We note in passing that using this to estimate

$|(V^n f)(t)|^2$ by means of the Cauchy-Schwarz inequality leads to $\|V^n\| \leq 1/n!$, so that V is quasinilpotent.

Recall that the *convolution of* functions $f, g \in L^1[0,b]$ is defined by

$$(f * g)(t) = \int_0^t f(t-s)g(s)ds \quad .$$

If e is the constantly 1 function on $[0,1]$, then $Vf = e * f$, so that $V^n f = e_n * f$, where e_n is the n-fold convolution of e with itself. The argument will be based on the following theorem of Titchmarsh [56] ([31] contains an interesting proof): if $f * g = 0$ a.e. in $[0,b]$ and $0 \in S(f)$, then $g = 0$ a.e. in $[0,b]$. Here the *support* $S(f)$ of f is the complement of the largest open set in which $f = 0$ a.e.

III. $f \in L^2[a,1]$ is a cyclic vector for $V \mid L^2[a,1]$ if and only if $a \in S(f)$.

Proof: Assume $g \in L^2[0,1]$ is such that $(V^n f, g) = 0$ for all $n \geq 0$. If $h(t) = \overline{g(1-t)}$, then

$$\begin{aligned}
(V^n f, g) &= \int_0^1 (V^n f)(t)h(1-t)dt \\
&= ((V^n f) * h)(1) \\
&= (e_n * f * h)(1) \\
&= (V^n(f * h))(1) \\
&= \frac{1}{(n-1)!} \int_0^1 (1-s)^{n-1}(f * h)(s)ds = \\
&= \frac{1}{(n-1)!} \int_0^1 t^{n-1}(f * h)(1-t)dt \quad ,
\end{aligned}$$

so that $f * h = 0$ a.e. in $[0,1]$. Now assume in addition that $f = 0$ in $[0, a]$ and that $a \in S(f)$. If f_1 is defined in $[0,1-a]$ by $f_1(t) = f(t+a)$, then

$$(f * h)(t+a) = (f_1 * h)(t)$$

in $[0,1-a]$, so that $f_1 * h = 0$ a.e. in $[0,1-a]$. But $0 \in S(f_1)$, and therefore $h = 0$ a.e. in $[0,1-a]$ by Titchmarsh's Theorem. Thus $g = 0$ a.e. in $[a,1]$, which proves that f is a cyclic vector for $V \mid L^2[a,1]$. The converse is clear.

IV. Theorem. $\mathfrak{M}$ is an invariant subspace of V if and only if $\mathfrak{M} = L^2[a,1]$ for some $a \in [0,1]$.

Proof: For $f \in L^2[0,1]$, let $a(f) = \inf S(f)$. Let $\mathfrak{M}$ be invariant. Then from III it follows that $L^2[a(f),1] \subset \mathfrak{M}$ for all $f \in \mathfrak{M}$. This implies that $L^2[a,1] \subset \mathfrak{M}$, where

$$a = \inf\{a(f) \mid f \in \mathfrak{M}\} \ .$$

Since the other inclusion is clear, $\mathfrak{M} = L^2[a,1]$ as required.

Consider now the *weighted shift operator* S, defined on ℓ_+^2 by

$$S(x_0, x_1, \ldots) = (0, \lambda_0 x_0, \lambda_1 x_1, \ldots) \ ,$$

where $\{\lambda_n\}$ is a bounded sequence. It is clear that for $n \geq 0$ the subspaces

$$\mathfrak{M}_n = \{x \mid x \in \ell_+^2 \text{ and } x_k = 0 \text{ for all } k < n\}$$

are invariant for S. Under suitable restrictions on the weights $\{\lambda_n\}$ it will be shown that these are the only nonzero invariant subspaces.

V. If $\lambda_n \to 0$ then S is quasinilpotent.

Proof: If $x = (x_0, x_1, \ldots)$, then $S^n x = y$, where $y_k = 0$ for $k < n$ and $y_{n+k} = \lambda_k \lambda_{k+1} \cdots \lambda_{k+n-1} x_k$ for $k \geq 0$. It follows that

$$\|S^n\| = \sup |\lambda_k \lambda_{k+1} \cdots \lambda_{k+n-1}| ,$$

and hence that $\|S^n\| \leq \mu_0 \mu_1 \cdots \mu_{n-1}$, where

$$\mu_k = \sup\{|\lambda_k|, |\lambda_{k+1}|, \ldots\} .$$

If $\lambda_n \to 0$ then $\mu_n \to 0$ and $(\mu_0 \mu_1 \cdots \mu_{n-1})^{1/n} \to 0$, so that S is quasinilpotent.

VI. Let $\mathfrak{M}$ be a subspace of ℓ^2_+. If there exists $\gamma \in \ell^2_+$ such that $|x_n| \leq \gamma_n \|x\|$ for all $x \in \mathfrak{M}$ and $n \geq 0$, then $\mathfrak{M}$ is finite-dimensional.

Proof: Let $e^1, \ldots, e^k$ be an orthonormal set in $\mathfrak{M}$, and let $f^n = (\overline{e}^1_n, \ldots, \overline{e}^k_n)$ for all $n \geq 0$. If a is any complex k-vector, then $x = \Sigma\, a_i e^i \in \mathfrak{M}$, and so

$$|(a, f^n)| = |\Sigma\, a_i e^i_n| = |x_n| \leq \gamma_n \|x\| = \gamma_n \|a\| .$$

This implies that

$$\sum_{i=1}^{k} |e^i_n|^2 = \|f^n\|^2 \leq \gamma_n^2$$

Summing over n gives $k \leq \|\gamma\|^2)$ so that $\dim \mathfrak{M} \leq \|\gamma\|^2 < \infty$.

VII. **Theorem.** (Nikolskii). Let S be the weighted shift operator with weights $\{\lambda_n\}$, and assume that $\{|\lambda_n|\}$ is non-increasing, $\lambda_n \neq 0$ for all $n \geq 0$, and $\Sigma\, |\lambda_n|^p < \infty$ for some $p \in (0, \infty)$.

Then $\mathfrak{M}$ is a nonzero invariant subspace of S if and only if $\mathfrak{M} = \mathfrak{M}_n$ for some $n \geq 0$.

Proof: Let n be the least integer for which there is $x \in \mathfrak{M}$ with $x_n \neq 0$. Then $\mathfrak{M} \subset \mathfrak{M}_n$, and it will be shown that equality holds. There is no loss of generality in assuming that $n = 0$, in which case it must be shown that $\mathfrak{M} = \ell^2_+$.

Fix an integer $N \geq 1$ such that $p \leq 2N$; then

$$\gamma = \{|\lambda_n|^N\} \in \ell^2_+ .$$

If $x \in \mathfrak{M}$ with $x_0 \neq 0$, then because $x, Sx, \ldots, S^N x$ are linearly independent elements of $\mathfrak{M}$, it is clear that a suitable linear combination is a vector $y \in \mathfrak{M}$ with $y_0 = 1$ and $y_1 = \cdots = y_N = 0$. Now let z be orthogonal to $\mathfrak{M}$, so that $(z, S^n y) = 0$ for all $n \geq 0$. Then

$$\sum_{k=0}^{\infty} \lambda_k \cdots \lambda_{k+n-1} y_k \bar{z}_{n+k} = 0$$

$$\bar{z}_n = -(\lambda_0 \cdots \lambda_{n-1})^{-1} \sum_{k=N+1}^{\infty} \lambda_k \cdots \lambda_{k+n-1} y_k \bar{z}_{n+k}$$

for all $n \geq 1$. But if $n, k \geq N + 1$,

$$\left|\frac{\lambda_k \cdots \lambda_{k+n-1}}{\lambda_0 \cdots \lambda_{n-1}}\right| \leq \left|\frac{\lambda_{N+1} \cdots \lambda_{N+n}}{\lambda_0 \cdots \lambda_{n-1}}\right|$$

$$= \left|\frac{\lambda_n \cdots \lambda_{N+n}}{\lambda_0 \cdots \lambda_N}\right| \leq \frac{|\lambda_n|^N}{|\lambda_0 \cdots \lambda_N|}$$

since $\{|\lambda_n|\}$ is non-increasing, and therefore

$$|z_n| \le B|\lambda_n|^N \|z\| \text{ for all } n \ge N+1 ,$$

where $B = |\lambda_0 \cdots \lambda_N|^{-1} \|y\|$. Hence $\dim \mathfrak{M}^\perp < \infty$ by VI. If $\mathfrak{M}^\perp \ne \{0\}$ there must be an eigenvector $e \in \mathfrak{M}^\perp$ for S^*. But S^* is quasinilpotent (because S is, by V), so $S^*e = 0$ and therefore $e_n = 0$ for $n \ge 1$. Since $(e, y) = 0$ and $y_0 = 1$ it follows that $e = 0$, a contradiction. Thus $\mathfrak{M}^\perp = \{0\}$ and $\mathfrak{M} = \ell^2_+$.

The rest of this section will be devoted to showing that operators that are close (in a suitable sense) to being unitary have nontrivial hyperinvariant subspaces. Call operators A and B *quasi-similar* if there are one-to-one operators P and Q each having dense range such that $AP = PB$ and $QA = BQ$. Similar operators have isomorphic lattices of invariant subspaces. Although this does not seem to be true for quasi-similarity, we have:

VIII. If A and B are quasi-similar and A has a nontrivial hyperinvariant subspace, then so does B.

Proof: There are one-to-one operators P and Q with dense range such that $AP = PB$ and $QA = BQ$. Suppose $\mathfrak{M}$ is hyperinvariant for A, and consider the subspace

$$\mathfrak{N} = \{ SQx \mid x \in \mathfrak{M} \text{ and } SB = BS \}^- .$$

It is clear that $\mathfrak{N}$ is hyperinvariant for B and that $\mathfrak{N} \ne \{0\}$ whenever $\mathfrak{M} \ne \{0\}$. If $SB = BS$ then

$$(PSQ)A = PSBQ = PBSQ = A(PSQ) ,$$

and therefore $PSQ\mathfrak{M} \subset \mathfrak{M}$ because $\mathfrak{M}$ is hyperinvariant for A. Hence $P\mathfrak{N} \subset \mathfrak{M}$, from which it follows that $\mathfrak{N} \neq \mathcal{H}$ whenever $\mathfrak{M} \neq \mathcal{H}$.

IX. Let T be a contraction such that $\|T^n x\| \not\to 0$ and $\|T^{*n}x\| \not\to 0$ for all $x \neq 0$. Then T is quasi-similar to a unitary operator.

Proof: As in the second proof of §6.I, let A be the non-negative square root of the strong limit of $T^{*n}T^n$, and let V be the isometry defined on $(A\mathcal{H})^-$ by $VA = AT$. Since $\|Ax\| = \lim \|T^n x\|$ the hypothesis implies that A is one-to-one, and hence that $(A\mathcal{H})^- = \mathcal{H}$. To complete the proof it will be shown that V is unitary and that there is a one-to-one operator B with dense range such that $BV = TB$. By hypothesis 0 is not an eigenvalue of T^*, and therefore T has dense range, V has dense range (from $VA = AT$), and V is unitary.

Let A_* and V_* be the operators constructed from T^* as above, so that $V_*A_* = A_*T^*$, $A_* = V_*^* A_* T^*$, and

$$A_* AVA = A_*A^2T = V_*^* A_*T^*A^2T = V_*^* A_*A^2 .$$

Since A has dense range this implies $A_*AV = V_*^* A_*A$. Now $B = A_*^2 A$ is one-to-one with dense range, and

$$\begin{aligned} TB &= TA_*^2 A = (A_*T^*)^*A_*A \\ &= A_*V_*^* A_*A = A_*^2 AV = BV \end{aligned}$$

as required.

X. Theorem. [54, Ch. II]. Let T be a contraction, and suppose that there are vectors x_0 and y_0 such that $\|T^n x_0\| \not\to 0$ and $\|T^{*n} y_0\| \not\to 0$. Then either T has a nontrivial hyperinvariant subspace or $T = cI$.

Proof: The subspaces $\{x \mid T^n x \to 0\}$ and $\{y \mid T^{*n} y \to 0\}^{\perp}$ are hyperinvariant for T. If both are proper then T is quasi-similar to a unitary operator V by IX. If V is scalar so is T; otherwise T has a nontrivial hyperinvariant subspace by VIII.

XI. Theorem. Let A be an operator such that $\operatorname{Re} A$ is of finite rank and $\operatorname{Re} A \leq 0$. Then A has a nontrivial invariant subspace.

Proof: According to §9 the Cayley transform

$$T = (A + I)(A - I)^{-1}$$

is an everywhere defined contraction. In addition, because

$$I - T^*T = -4(A^* - I)^{-1}(\operatorname{Re} A)(A - I)^{-1} ,$$

$I - T^*T$ is of finite rank. In the same way so is $I - TT^*$. If $T^n \to 0$ strongly, then by §2.II T is unitarily equivalent to a part of a backward shift of finite multiplicity. By a result mentioned at the end of §2 (and also at the end of §4), T has a nontrivial invariant subspace. If $T^{*n} \to 0$ strongly, then in the same way T^* has a nontrivial invariant subspace, and hence so does T. If neither of these is the case, a nontrivial invariant subspace exists by X. Thus in all cases T has a nontrivial invariant subspace. The proof is completed by showing $T\mathfrak{M} \subset \mathfrak{M}$ implies $A\mathfrak{M} \subset \mathfrak{M}$. Since

$$T = I + 2(A-I)^{-1},$$

$T\mathfrak{M} \subset \mathfrak{M}$ implies $(A-I)^{-1}\mathfrak{M} \subset \mathfrak{M}$. If the inclusion were proper, there would exist $x \neq 0$ such that $((A-I)^{-1}x, x) = 0$, so that for $y = (A-I)^{-1}x$,

$$0 = (y, (A-I)y) = (y, Ay) - \|y\|^2 \leq -\|y\|^2,$$

and consequently $y = 0$ and $x = 0$, a contradiction. Hence $(A-I)^{-1}\mathfrak{M} = \mathfrak{M}$, $(A-I)\mathfrak{M} = \mathfrak{M}$, and $A\mathfrak{M} \subset \mathfrak{M}$.

REMARK. This result is valid if $\operatorname{Re} A$ is compact and its sequence of eigenvalues lies in ℓ_+^p for some $p \in [1, \infty)$ [46]. The question is open if $\operatorname{Re} A$ is merely compact.

Exercises. 1. If p is a polynomial, the null space and the closure of the range of $p(T)$ are hyperinvariant subspaces for T. When $\mathcal{H}$ is finite-dimensional every hyperinvariant subspace is of this form.

2. If $\mathfrak{M}$ is hyperinvariant for T then $\mathfrak{M}^\perp$ is hyperinvariant for T^*.

3. Quasi-similarity is an equivalence relation. If A and B are quasi-similar, so are A^* and B^*.

4. Under the hypotheses of IX, the operators V and V_*^* are unitarily equivalent.

5. Theorem VII is valid for weighted shift operators on ℓ_+^q, $1 \leq q < \infty$. (Prove the following and use it to establish the required generalization of VI: if $\{T_n\}$ is a sequence of operators on a Banach space B such that $T_n \to I$ strongly, then a subset C of B is relatively compact if and only if $T_n \to I$ uniformly on C.)

SECTION 11.

INVARIANT SUBSPACES FOR COMPACT OPERATORS

In this section use will be made of the weak operator topology on the algebra $\mathcal{B}(\mathcal{H})$ of all operators on $\mathcal{H}$. A net $\{T_\alpha\}$ converges to an operator T in this topology if and only if $(T_\alpha x, y) \to (Tx, y)$ for all $x, y \in \mathcal{H}$. A typical neighborhood U of T is determined by vectors $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathcal{H}$ as follows:

$$U = \{S \mid |((T-S)x_i, y_i)| \leq 1 \text{ for } 1 \leq i \leq n\} .$$

1. Theorem. The unit ball of $\mathcal{B}(\mathcal{H})$ is compact in the weak operator topology.

Proof: Since $\mathcal{H}$ is reflexive, the ball

$$B_r = \{x \mid x \in \mathcal{H} \text{ and } \|x\| \leq r\}, \quad r \geq 0$$

is a weakly compact subset of $\mathcal{H}$ by Alaoglu's Theorem. Therefore $\wp = \Pi\{B_{\|x\|} \mid x \in \mathcal{H}\}$ is compact in the cartesian product topology. This product consists of the functions $f\colon \mathcal{H} \to \mathcal{H}$ such that $\|f(x)\| \leq \|x\|$ for all $x \in \mathcal{H}$, and consequently contains the unit ball of $\mathcal{B}(\mathcal{H})$. It is almost obvious that the unit ball of $\mathcal{B}(\mathcal{H})$ is a closed subset of $\wp$ on which the product topology and the weak operator topology coincide, and the theorem follows.

In outline, the method for obtaining invariant subspaces is to produce projections E_n whose ranges are almost invariant, in a suitable sense, and then to consider a cluster point of $\{E_n\}$ in the weak operator topology (which must exist by I). Unfortunately such a cluster point need not be a projection. In fact:

II. Theorem. If $\mathcal{H}$ is infinite-dimensional, the closure in the weak operator topology of the set $\mathcal{E}$ of all projections is the set $\mathcal{C}^+$ of all positive contractions.

LEMMA. If A is a positive contraction and Q is a projection such that $\dim (I-Q)\mathcal{H} = \dim \mathcal{H}$, there is a projection E such that $QAQ = QEQ$.

Proof: By §7 there is a space $\mathcal{K} \supset \mathcal{H}$ and a projection F in $\mathcal{K}$ with $PF\,|\,\mathcal{H} = A$ and $\dim \mathcal{K} = \dim \mathcal{H}$ (where P is the projection on $\mathcal{H}$). From this and the hypothesis it follows that that $\dim(\mathcal{K} \ominus \mathfrak{D}) = \dim(\mathcal{H} \ominus \mathfrak{D})$, where $\mathfrak{D} = Q\mathcal{H}$. Therefore there is an isometry W from $\mathcal{H}$ onto $\mathcal{K}$ such that $W\,|\,\mathfrak{D}$ is the identity. If $E = W^*FW$, then for all $x, y \in \mathcal{H}$,

$$(QEQx, y) = (W^*FWQx, Qy) = (FWQx, WQy) = (FQx, Qy)$$

$$= (AQx, Qy) = (QAQx, y) \quad .$$

Proof of Theorem. It is clear that $\mathcal{C}^+$ contains the closure of $\mathcal{E}$. Consider a positive contraction A. Write

$$\mathcal{H} = \sum_{k \geq 1} \oplus \mathcal{H}_k$$

with $\dim \mathcal{H}_k = \dim \mathcal{H}$ for all $k \geq 1$, and let Q_n be the projection on $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$. By the lemma there is a projection

E_n such that $Q_n E_n Q_n = Q_n A Q_n$ for all $n \geq 1$. Now $Q_n \to I$ strongly, and it follows that $E_n \to A$ weakly, for

$$((E_n - A)x, y) = ((E_n - A) Q_n x, Q_n y) + ((E_n - A) Q_n x, (I - Q_n)y)$$
$$+ ((E_n - A)(I - Q_n)x, y)$$

and $((E_n - A) Q_n x, Q_n y) = 0$, so that

$$|((E_n - A)x, y)| \leq 2\|x\| \, \|(I - Q_n)y\| + 2\|(I - Q_n)x\| \, \|y\| \, .$$

This result is due to Halmos [20] and Nagy [49]. By arguing in the same way with unitary dilations, it may be shown that the closure in the weak operator topology of the unitary operators is the entire unit ball of $\mathcal{B}(\mathcal{H})$.

The construction of invariant subspaces begins with the following lemma.

III. Let $e \in \mathcal{H}$, $T \in \mathcal{B}(\mathcal{H})$, R_n the projection on $[e, Te, \ldots, T^n e]$, and $d_n = \|T^n e - R_{n-1} T^n e\|$ the distance from $T^n e$ to $[e, \ldots, T^{n-1} e]$. Then $\|TR_n - R_n T R_n\| = d_{n+1}/d_n$.

Proof: Fix $x \in \mathcal{H}$ and let $R_n x = a_0 e + \cdots + a_n T^n e$. Then

$$TR_n x = a_0 Te + \cdots + a_n T^{n+1} e$$

$$R_n T R_n x = a_0 Te + \cdots + a_{n-1} T^n e + a_n R_n T^{n+1} e$$

$$TR_n x - R_n T R_n x = a_n (T^{n+1} e - R_n T^{n+1} e)$$

$$\|TR_n x - R_n T R_n x\| = |a_n| d_{n+1} = (d_{n+1}/d_n)(d_n |a_n|).$$

But
$$|a_n| d_n = \|a_n T^n e - R_{n-1}(a_n T^n e)\|$$
$$= \| R_n x - R_{n-1} R_n x\|$$
$$= \|R_n x - R_{n-1} x\|,$$

and the lemma follows.

IV. If e is a cyclic vector for T such that $\liminf \|T^n e\|^{1/n} = 0$, then there are finite-dimensional projections $P_1 \leq P_2 \leq \cdots$ such that $P_n \to I$ strongly and $\|TP_n - P_n TP_n\| \to 0$. If Q_n is a projection such that $Q_n \leq P_n$ and $Q_n \mathcal{H}$ is invariant for $T_n = P_n T P_n$, then $\|Q_n T Q_n - TQ_n\| \to 0$.

Proof: $\liminf (d_{n+1}/d_n) \leq \liminf d_n^{1/n} \leq \liminf \|T_n e\|^{1/n} = 0$. Therefore if $\{P_n\}$ is a suitable subsequence of $\{R_n\}$, III implies $\|TP_n - P_n TP_n\| \to 0$. Since e is cyclic $R_n \to I$, so $P_n \to I$ too. Finally

$$TQ_n - Q_n TQ_n = TQ_n - Q_n P_n TP_n Q_n$$
$$= TQ_n - Q_n T_n Q_n$$
$$= TQ_n - T_n Q_n$$
$$= (TP_n - P_n TP_n) Q_n ,$$

so $\|TQ_n - Q_n TQ_n\| \leq \|TP_n - P_n TP_n\| \to 0$.

The next step is to construct projections Q_n as in IV, but with the additional property that if Q is a weak cluster point of $\{Q_n\}$, then $Q \neq 0, I$. This will be done with the help of the linear functional

$$\rho(A) = \tfrac{1}{2}[(Ae, e) + (Af, f)],$$

where e and f are any orthogonal unit vectors. It is easy to see that $\rho(E) \leq \frac{1}{2}$ for any one-dimensional projection E. Since $\rho(P_n) \to 1$, it can be assumed that $\rho(P_n) \geq 3/4$ for all n. If P_n is m-dimensional, there are projections $P_n^1 \leq P_n^2 \leq \cdots \leq P_n^m = P_n$ such that $\dim P_n^k = k$ and $P_n^k \mathcal{H}$ is invariant for $P_n T P_n$, $k = 1, 2, \ldots, m$. Since

$$\rho(P_n^{k+1}) - \rho(P_n^k) = \rho(P_n^{k+1} - P_n^k) \leq \frac{1}{2} \quad ,$$

there is $j \leq m$ with $\frac{1}{4} \leq \rho(P_n^j) \leq \frac{3}{4}$. Let $Q_n = P_n^j$. If Q is a weak cluster point of $\{Q_n\}$, then $\frac{1}{4} \leq \rho(Q) \leq \frac{3}{4}$, and therefore $Q \neq O, I$.

V. The space $\mathfrak{M} = \{x \mid Qx = x\}$ is closed, invariant, and distinct from $\mathcal{H}$.

Proof: Obviously $\mathfrak{M}$ is a closed subspace, and $\mathfrak{M} \neq \mathcal{H}$ since $Q \neq I$. Let $\{Q_a\}$ be a subnet of $\{Q_n\}$ which converges weakly to Q. If $x \in \mathfrak{M}$, then

$$\|Q_a x - x\|^2 = (x, x) - (Q_a x, x) \to 0 \quad ,$$

and hence $\|TQ_a x - Tx\| \to 0$. But

$$\begin{aligned} |(Q_a T Q_a x - QTx, y)| &\leq |(Q_a T Q_a x - Q_a Tx, y)| + |(Q_a Tx - QTx, y)| \\ &\leq \|TQ_a x - Tx\| \, \|y\| + |((Q_a - Q)Tx, y)| \quad , \end{aligned}$$

so that $Q_a T Q_a x \to QTx$ weakly. On the other hand, $Q_a T Q_a x \to TQx$ weakly by IV, so that $QTx = TQx = Tx$ and $Tx \in \mathfrak{M}$.

The only point remaining is whether $\mathfrak{M} \neq \{0\}$, and this is where a compactness hypothesis is used.

VI. Theorem. Let T be an operator such that (i) there is a non-zero vector e with $\liminf \|T^n e\|^{1/n} = 0$, and (ii) the norm-closed algebra generated by T and I contains a non-zero compact operator C. Then T has a nontrivial invariant subspace.

Proof: Since the subspace spanned by $e, Te, \ldots$ is invariant, it can be assumed that e is a cyclic vector. Consider the set $\mathfrak{A}$ of all operators A such that $\|Q_a A Q_a - A Q_a\| \to 0$, where $\{Q_a\}$ is as above. For any operators A and B,

$$\begin{aligned}\|Q_a B Q_a - B Q_a\| &\leq \|Q_a B Q_a - Q_a A Q_a\| + \|Q_a A Q_a - A Q_a\| \\ &\quad + \|A Q_a - B Q_a\| \\ &\leq \|Q_a A Q_a - A Q_a\| + 2\|A - B\| ,\end{aligned}$$

and consequently $\mathfrak{A}$ is norm-closed. Clearly $\mathfrak{A}$ is linear, and since

$$\begin{aligned}Q_a(AB)Q_a - (AB)Q_a &= (Q_a A Q_a - A Q_a) B Q_a \\ &\quad + (A - Q_a A)(Q_a B Q_a - B Q_a) ,\end{aligned}$$

it is an algebra. But $T \in \mathfrak{A}$ by IV, and so $C \in \mathfrak{A}$:

$$\|Q_a C Q_a - C Q_a\| \to 0 .$$

From this and the compactness of C it follows readily that $QCQ = CQ$, and hence that $CQ\mathcal{H} \subset \mathfrak{M}$. For $CQ_a \to CQ$ strongly since C is compact, and therefore

$$Q_a C Q_a - QCQ = Q_a(CQ_a - CQ) + (Q_a - Q)CQ ,$$

$$|((Q_\alpha C Q_\alpha - QCQ)x, y)| \le \|(CQ_\alpha - CQ)x\| \; \|y\| + |((Q_\alpha - Q)CQx, y)| \quad ,$$

and $Q_\alpha C Q_\alpha \to QCQ$ weakly.

If $\mathfrak{M} \neq \{0\}$ there is nothing to prove. If $\mathfrak{M} = \{0\}$, then $CQ\mathcal{H} = \{0\}$, so that the null space of C is nontrivial (because $Q \neq 0$). But C and T commute, and therefore the null space of C is invariant for T. This completes the proof.

COROLLARY. If T is quasinilpotent and if the norm-closed algebra generated by T and I contains a non-zero compact operator, then T has a nontrivial invariant subspace.

These results are due to Arveson and Feldman [2]; the proof is based on earlier work of Bernstein and Robinson [7] and Halmos [23].

VII. Theorem. If $p(T)$ is compact for some polynomial $p \neq 0$, then T has a nontrivial invariant subspace.

Proof: Let $C = p(T)$. If $0 \neq \lambda \in \sigma(C)$ and γ is a contour enclosing λ but no other point of $\sigma(C)$, then

$$E = \frac{1}{2\pi i} \int_\gamma (zI - C)^{-1} dz$$

is an idempotent distinct from 0 and I which commutes with every operator commuting with C. In particular E and T commute, so $E\mathcal{H}$ is a nontrivial invariant subspace.

On the other hand, suppose that $\sigma(C) = \{0\}$. In this case $\sigma(T)$ is a finite set by the Spectral Mapping Theorem. If there is more than one point in this set, then the above argument may be repeated. If $\sigma(T) = \{\lambda\}$, then $T - \lambda I$ is quasinilpotent and VI applies, provided $C \neq 0$. Since it is obvious that T has invariant subspaces when $p(T) = C = 0$, the theorem is proved.

COROLLARY. Any compact operator has a nontrivial invariant subspace.

The theorem is due to Bernstein and Robinson [7], and the corollary to von Neumann, Aronszajn, and Smith [1] (on any Banach space).

REFERENCES

1. N. Aronszajn and K. T. Smith, "Invariant subspaces of completely continuous operators," *Ann. Math* 60 (1954), 345-350.
2. W. B. Arveson and J. Feldman, "A note on invariant subspaces," *Mich. Math. J.,* to appear.
3. S. K. Berberian, *Introduction to Hilbert Space,* Oxford University Press, New York, 1961.
4. ——, *Notes on Spectral Theory,* Van Nostrand, Princeton, 1966.
5. C. A. Berger, "A strange dilation theorem, " *Notices A.M.S.* 12 (1965), 590.
6. C. A. Berger and J. G. Stampfli, "Mapping theorems for the numerical range, *Am. J. Math.* 89 (1967), 1047-1055.
7. A. R. Bernstein and A. Robinson, "Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos," *Pac. Math. J.* 16 (1966), 421-431.
8. A. Beurling, "On two problems concerning linear transformations in Hilbert space," *Acta. Math.* 81 (1949), 239-255.
9. J. Bram, "Subnormal operators," *Duke Math. J.* 22 (1955), 75-94.
10. L. de Branges and J. Rovnyak, *Square Summable Power Series,* Holt, Rinehart, and Winston, 1966.

11. ——, *Perturbation Theory and Its Applications in Quantum Mechanics* (Calvin H. Wilcox, editor), Wiley, 1966.

12. A. Brown, "On a class of operators," *Proc. A.M.S.* 4 (1953), 723-728.

13. —— and P. R. Halmos, "Algebraic properties of Toeplitz operators," *J. Reine Angew. Math.* 213 (1963), 89-102.

14. J. L. B. Cooper, "One parameter semi-groups of isometric operators in Hilbert space," *Ann. Math.* (2) 48 (1947), 827-842.

15. R. G. Douglas, "Structure theory for operators, I," *J. Reine Angew. Math.*, 232 (1968), 180-193.

16. N. Dunford and J. Schwartz, *Linear Operators, Part I*, Interscience, New York, 1958.

17. P. A. Fillmore and D. M. Topping, "A direct integral decomposition for certain operator algebras," *Am. J. Math.*, 41 (1969), 11-17.

18. S. R. Foguel, "Powers of a contraction in Hilbert space," *Pac. J. Math.* 13 (1963), 551-562.

19. R. Goodman, "Invariant subspaces for normal operators," *J. Math. Mech.* 15 (1966), 123-128.

20. P. R. Halmos, "Normal dilations and extensions of operators, *Summa Brasil.* 2 (1950), 125-134.

21. ——, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea, New York, 1957.

22. ——, "Shifts on Hilbert spaces," *J. Reine Angew. Math.* 208 (1961), 102-112.

23. ——, "Invariant subspaces for polynomially compact operators," *Pac. J. Math.* 16 (1966), 433-437.

24. ——, *A Hilbert Space Problem Book,* Van Nostrand, Princeton, 1967.

25. I. Halperin, "The unitary dilation of a contraction operator," *Duke Math. J.* 28 (1961), 563-572.

26. ——, "The product of projection operators," *Acta. Sci. Math.* (Szeged) 23 (1962), 96-99.

27. F. Hartogs and A. Rosenthal, "Uber Folgen Analytischer Functionen," *Math. Ann.* 104 (1930-1931).

28. H. Helson, *Invariant Subspaces,* Academic Press, New York, 1964.

29. E. Hille and R. S. Phillips, *Functional Analysis and Semigroups,* American Mathematical Society, Providence, 1957.

30. K. Hoffman, *Banach Spaces of Analytic Functions,* Prentice-Hall, Englewood Cliffs, N.J., 1962.

31. P. D. Lax, "Translation invariant subspaces," *Acta. Math.* 101 (1959), 163-178.

32. —— and R. S. Phillips, *Scattering Theory,* Academic Press, 1967.

33. P. Masani, "Isometric flows on Hilbert space," *Bull. Amer. Math. Soc.* 68 (1962), 624-632.

34. M. A. Naimark, "Positive definite operator functions on a commutative group," *Izv. Akad. Nauk SSSR Ser. Mat.* 7 (1943), 237-244.

35. ——, "On a representation of additive operator set functions," *Comptes Rendus (Doklady) Acad. Sci. URSS* 41 (1943), 359-361.

36. ——, *Normed Rings,* Noordhoff, 1964.

37. J. von Neumann, "Allgemeine Eigenwerttheorie Hermitescher Functionaloperatoren," *Math. Ann* 102 (1929), 49-131.

38. ——, "Eine Spektraltheorie fur allgemeine Operatoren eines unitären Raumes," *Math. Nachrichten* 4 (1951), 258-281.

39. V. P. Potapov, "The multiplicative structure of J-contractive matrix functions," *Trudy Moscov. Mat. Obsc.* 4 (1955), 125-236. (Russian); *A.M.S. Transl.* 15 (1960), 131-243.

40. P. Rosenthal, "A note on unicellular operators," *Proc. Amer. Math. Soc.* 19 (1968), 505-506.

41. G.-C. Rota, "On models for linear operators," *Comm. Pure Appl. Math.* 13 (1960), 469-472.

42. W. Rudin, *Fourier Analysis on Groups,* Interscience, New York, 1962.

43. D. Sarason, "A remark on the Volterra operator," *J. Math. Anal. Appl.* 12 (1965), 244-246.

44. J. J. Schäffer, "On unitary dilations of contractions," *Proc. A.M.S.* 6 (1955), 322.

45. M. Schreiber, "Remark on a paper of Kalisch," *J. Math. Anal. Appl.* 7 (1963), 62-63.

46. J. Schwartz, "Subdiagonalization of operators in Hilbert space with compact imaginary part," *Comm. Pure Appl. Math.* 15 (1962), 159-172.

47. J. G. Stampfli, "Hyponormal operators," *Pac. J. Math.* 12 (1962), 1453-1458.

48. B. Sz.-Nagy, "Transformations de l'espace de Hilbert, fonctions de type positif sur un groupe," *Acta. Sci. Math. (Szeged)* 15 (1954), 104-114.

49. ——, "Suites faiblement convergentes de transformations normales de l'espace Hilbertien," *Acta. Math. Acad. Sci. Hung.* 8 (1957), 295-301.

50. ——, "Isometric flows in Hilbert space," *Proc. Camb. Phil. Soc.* 60 (1964), 45-49.

51. ——, Appendix to *Functional Analysis* by F. Riesz and B. Sz.-Nagy, Ungar, New York, 1960.

52. B. Sz.-Nagy and C. Foias, "Sur les contractions de l'espace de Hilbert IV," *Acta. Sci. Math. (Szeged)* 21 (1960), 251-259.

53. ——, "Sur les contractions de l'espace de Hilbert VIII," *Acta. Sci. Math. (Szeged)* 25 (1964), 38-71.

54. ——, *Analyse Harmonique des Operateurs de l'Espace de Hilbert,* Akademiai Kiado, Budapest, 1967.

55. J. L. Taylor, "The Tomita decomposition of rings of operators," *Trans. A.M.S.* 113 (1964), 30-39.

56. E. C. Titchmarsh, "The zeros of certain integral functions," *Proc. Lond. Math. Soc. (2)* 25 (1926), 283-302.

57. J. Wermer, "On invariant subspaces of normal operators," *Proc. A.M.S.* 3 (1952), 270-277.

58. ——, "Report on Subnormal Operators, Operator Theory, and Group Representation," *NAS-NRC,* 1953.

59. J. P. Williams, "Similarity and the numerical range," *J. Math. Anal. Appl.,* 26 (1969), 307-314.

INDEX

VAN NOSTRAND REINHOLD MATHEMATICAL STUDIES are paperbacks focusing on the living and growing aspects of mathematics. They are not reprints, but original publications. They are intended to provide a setting for experimental, heuristic and informal writing in mathematics that may be research or may be exposition.

Under the editorship of Paul R. Halmos and Frederick W. Gehring, lecture notes, trial manuscripts, and other informal mathematical studies will be published in inexpensive, paperback format.

P. R. HALMOS received his Ph.D. from the University of Illinois, and spent three years at the Institute for Advanced Study, two them as Assistant to John von Neumann. He taught at the versities of Chicago, Michigan, and Hawaii and is presen fessor of Mathematics at Indiana University.

F. W. GEHRING received his Ph.D. from Camb England. He has held Visiting Professorships at ford Universities as well as Guggenheim. Fellowships at the University of Helsink Technische Hochschule in Zürich. H Mathematics at the University of Mi

C2390

A VAN NOSTRAND REINHOLD MATHEMATICAL STUDY

•

About this book:

The notes developed from a series of lectures given at Indiana University. The audience was familiar with the basic theory of Hilbert spaces and operators up through the spectral theory. This book is intended to be an introduction to structure theory for non-normal bounded linear operators on Hilbert space, and consequently attempts to provide a reasonably broad background.

Although material is drawn from a wide variety of sources, the point of view is usually that developed by Sz-Nagy, Foias, and others which relates the structure of an operator to the behavior of the sequence of iterates. There is a through treatment of shift operators and their invariant subspaces, leading to a development of model theory and dilation theory for a single operator. These ideas are carried further in extensive sections dealing with one-parameter contractive semigroups and with the circle of results in operator-valued mappings. Additional topics include hypernormal operators, weighted shift operator variant subspaces, and invariant s

•

The Author:

PETER A. FILLMORE (Ph.D. University of Minnesota) is currently Associate Professor of Mathematics at Indiana University. Prior to this he was an instructor at the University of Chicago. He has contributed extensively to the journals in the area of operator theory.

VAN NOSTRAND REINHOLD COMPANY
450 West 33rd Street New York, N.Y. 10001