SPECTRAL THEORY FOR A CLASS
OF NON-NORMAL OPERATORS

HARRY GONSHOR

1. Introduction. As is well known, the spectral theorem plays an important part in mathematics because of its many applications. Unfortunately, the theorem is valid for normal operators only. In view of this, attempts have been made by several mathematicians to obtain a theorem about a more general class of operators, which will reduce to the ordinary spectral theorem if the operator is normal. Brown **(1)** has developed a unitary equivalence theory for a certain class of operators. The present paper builds up a spectral theory for a class of operators properly containing the set of all normal operators. The chief technique used is that of direct integral theory. For many reasons the results as well as the methods of proof seem natural to the author.

We shall assume an elementary knowledge of Banach Algebra theory, spectral theory, and direct integral theory. All spaces will be assumed to be separable.

2. J_n operators. We begin with a definition: A is a J_n operator if and only if there exists a direct integral decomposition of the space H such that A is decomposable and such that if $A(t)$ is expressed in matrix form for each point t in the space with respect to which H is decomposed, it is of the form

$$D_m = \begin{pmatrix} D & 0 & 0 \dots \\ 0 & D & 0 \dots \\ 0 & 0 & D \dots \\ \ddots & \ddots & \ddots \end{pmatrix}$$

where the order of D is less then or equal to n. We shall use the term "pure J_n operator" to refer to the case where we insist that the order of D is exactly n. In the sequel we shall let D_m be an m-fold copy of D (m possibly infinite).

Note that A is a J_1 operator if and only if A is normal. This is a consequence of the definition and of the spectral theorem. We also make the trivial remark that an nth order matrix is a J_n operator.

We shall assume that $a_1, a_2, a_3, \dots, \in H$ are the basic elements used. This means that $\{a_i(t)\}$ is an orthogonal basis for $H(t)$ for all t where $H(t)$ is the Hilbert space at the point t, and i runs from 1 to ∞ if dim $H(t) = \infty$ and from 1 to n if dim $H(t) = n$. It will also be assumed that this is the basis used in order to obtain the matrix form for $A(t)$.

Now let A be a J_n operator. For each t let $r(t)$ be the least r such that $A(t)$ is of the form D_m where D has order r. (By definition, $r \leqslant n$.) We assert that

$r(t)$ is measurable as a function of t. To begin with, the set of all t where dim $H(t)$ is a multiple of r_0 is measurable. Consider the equations

$$[A(t) \, a_m(t), \quad a_n(t)] = 0.$$

where $m < a \leqslant n$ or $n < a \leqslant m$ for some $a \equiv 1 \pmod{r_0}$. This is only a countable number of conditions, and hence the set on which all the conditions are simultaneously satisfied is measurable. If it is also demanded that the conditions are not simultaneously satisfied for any integer less than r_0, then the set which satisfies all the requirements is still measurable. But the requirements are nothing but a restatement of the condition that $r(t) = r_0$. Hence $r(t)$ is measurable. This shows that a J_n operator can be expressed as a direct sum of pure J_m operators. We can therefore restrict ourselves to pure J_n operators.

Let H be decomposed into spaces $H(t)$ all of which have dimension divisible by n. (For this purpose we adopt the convention that ∞ is divisible by n.) We shall consider operator functions $A(t)$ which are of the form $D_m(t)$ where the order of D_m is n for all t. Let D be written explicitly as

$$\begin{pmatrix} a_{11}(t) \, \ldots \, a_{1n}(t) \\ \ldots \\ a_{n1}(t) \, \ldots \, a_{nn}(t) \end{pmatrix}$$

LEMMA 2.1. $A(t)$ is an operator $\rightleftarrows a_{ij}(t) \in L^\infty$ for all i and j.

The proof is trivial and is therefore left to the reader. We remind the reader that a necessary and sufficient condition for an operator function $A(t)$ to represent an operator A is that $[A(t) \, a_m(t), a_n(t)]$ be measurable for all m and n and that $||A(t)|| \in L^\infty$.

THEOREM 1. A is a pure J_n operator $\rightleftarrows H$ can be decomposed into n mutually orthogonal equivalent projections E_i (also called U_{ii}) with partial isometries U_{ij} satisfying $U_{ij} \, U_{kl} = \delta_{jk} \, U_{il}$ such that $A = \sum_{i,j} A_{ij} \, U_{ij}$ where the A_{ij} are mutually commuting normal operators which commute with the U_{kl}.

Proof. Define $E_i(t)$ to be the projection on $[a_i(t), a_{i+n}(t), \ldots a_{i+dn}(t) \ldots]$ for all i satisfying $1 \leqslant i \leqslant n$. Clearly $E_i(t)$ is an operator. Define $U_{ij}(t)$ as the partial isometry from $E_j(t)$ to $E_i(t)$ which maps $a_{j+dn}(t)$ into $a_{i+dn}(t)$ for all integers d and all other $a_k(t)$ into zero. It can immediately be checked that $U_{ij}(t)$ is measurable and bounded, and hence that it defines an operator on H. Since the property of being partially isometric is purely algebraic, equations in $H(t)$ carry over into equations on H. Note that

$$U_{ij} \, U_{kl} = \delta_{jk} \, U_{il}.$$

Also the $\{E_i\}$ are mutually orthogonal projections with sum 1. Now

$$A(t) = \sum_{i,j} \lambda_{ij}(t) \, U_{ij}(t)$$

where $\lambda_{ij}(t)$ are scalar functions. Hence

$$A = \sum_{i,j} A_{ij} U_{ij}$$

where the A_{ij} satisfy the conditions in the right-hand side in Theorem 1.

Now suppose A satisfies the condition in the right-hand side. It is clear that $R(A_{ij}, A_{ij}{}^*)$ is an Abelian ring commuting with the U_{ij}, hence in particular with the E_i. We shall decompose with respect to R. Now R can be regarded as an Abelian ring of operators on the space E_1. We can therefore choose a set of basic elements for a decomposition of E_1 which we call f_1, $f_{n+1}, f_{2n+1} \ldots f_{dn+1} \ldots$ (where d is an integer). We define f_{dn+j} where $1 < j \leqslant n$ as $U_{ji} f_{dn+1}$. It is easy to check that $f_1, f_2 \ldots$ can be used as a set of basic elements for a decomposition of H in which R is precisely the set of all operators which become scalar operator functions. By construction of the f's it is easily seen that $U_{ij}(t)$ is nothing but the operator which, when expressed in matrix form, has 1's in the $(dn + i)$th row and $(dn + j)$th column for all d and 0's everywhere else. It follows that if A_{ij} is mapped into the scalar function $a_{ij}(t)$,

$$A = \sum_{i,j} A_{ij} U_{ij}$$

is mapped into $D_m(t)$ where

$$D(t) = \begin{pmatrix} a_{11}(t) \ldots a_{1n}(t) \\ \phantom{a_{11}(t)} \cdots \\ a_{n1}(t) \ldots a_{nn}(t) \end{pmatrix}.$$

This completes the proof of the Theorem.

Theorem 1 gives an intrinsic characterization of J_n operators in the large, that is, no direct integral theory is required in the definition. Alternative characterizations of J_n operators are as follows:

(1) $R(A, A^*)$ when decomposed with respect to its center gives rise to factors of type I_m only where $m \leqslant n$.

(2) $R(A, A^*)$ has at most n orthogonal equivalent non-zero projections.

3. Existence of spectral representations. We begin with certain results regarding measurability.

THEOREM 2. *Let T be a measure space, and let $f_t(x)$ be nth degree polynomials $a_0(t) x^n + a_1(t) x^{n-1} \ldots + a_n(t)$ which are defined for all $t \in T$ and such that $a_i(t)$ are measurable for all i. Then there exists a measurable function $y(t)$ such that $y(t)$ is a root of $f_t(x) = 0$ for all t.*

For this purpose we shall choose a complete ordering of the complex numbers which will be used throughout the remainder of the paper. We let 0 be the least and express any other complex number in the form $re^{i\theta}$ where $0 \leqslant \theta < 2\pi$, ordering them lexicographically with respect to r and θ. Theorem 2 will be

shown by letting $y(t)$ be the least root of $f_t(x)$ for each t. We use several lemmas.

LEMMA 3.1. *Let k be a compact subset of the complex plane. Then the set of all t such that $f_t(z)$ has a zero in k is measurable.*

Proof. For any fixed z_0,

$$a_0(t)z_0{}^n + a_1(t)\, z_0{}^{n-1} + \ldots + a_n(t)$$

is a linear combination of measurable functions and is hence measurable. Now let z run through all points of a given countable dense subset of k. We conclude that [g.l.b. $|f_t(z)|$] is measurable, hence the set where [g.l.b. $|f_t(z)|$] is 0 is measurable. Since k is compact, this condition is equivalent to the condition that $f_t(z)$ has a zero in k. This proves the lemma.

In practice the compact sets used will be one of two special types: the closed region bounded by the circle $r = r_0$, and the "conical" region bounded by the circle $r = r_0$ and the lines $\theta = 0$ and $\theta = \theta_0$.

LEMMA 3.2. *The minimum modulus of the roots of $f_t(x)$ is a measurable function with respect to t.*

Proof. It is required to prove that the set where the minimum modulus is above r_0 is measurable. But the minimum modulus is above r_0 if and only if the region bounded by $r = r_0$ has no zeros. Thus the result follows from Lemma 4.1.

LEMMA 3.3. *For every t take the least argument $\theta(t)$ of all roots of $f_t(x)$ of minimum modulus. Then $\theta(t)$ is a measurable function of t. It is understood that the argument is taken to be non-negative and less than 2π.*

Proof. We shall prove that the set where $\theta(t)$ exceeds θ_0 is measurable. Let s_r be the closed region bounded by the circle with center at the origin and radius r, and let C_r be the intersection of S_r and the angular region $0 \leqslant \theta \leqslant \theta_0$. Let r run through all the rationals. We now consider the statement that for some r, S_r has a zero of $f_t(x)$ but C_r does not. Since the set where S_r has a zero is measurable with respect to t and similarly for C_r, the set where the above statement is true is measurable. (Note that only rational radii were used.) It remains to show that the above statement is equivalent to the fact that $\theta(t) > \theta_0$.

Suppose that for some r, S_r has a zero but C_r does not. Then the minimum modulus of the roots does not exceed r. Since C_r has no zero, every root of minimum modulus has argument greater than θ_0. Thus $\theta(t) > \theta_0$.

On the other hand, suppose $\theta(t) > \theta_0$. If r is the minimum modulus of the roots, then S_r has a zero and C_r does not. To complete the proof we need a rational r having this property. Now a polynomial has only a finite number of roots. Hence there is a least modulus $r_1 > r$ which the roots can have. Let d be a rational strictly between r and r_1. Clearly S_d has a zero, because even S_r

does. C_d has no zeros because C_r does not; and by choice of d, $C_d - C_r$ has no zeros. This completes the proof of the lemma.

Theorem 2 now follows trivially from Lemmas 3.2 and 3.3.

COROLLARY. *It is possible to select n measurable functions $y_i(t)$ such that for each t, $\{y_i(t)\}$ are the roots of $f_t(x)$ appearing with their proper multiplicity. In fact, the functions may be chosen so that $y_i(t) \leqslant y_{i+1}(t)$.*

Proof. It suffices to remark that the coefficients of $f_t(y)/\{y - y_1(t)\}$ are measurable because of the nature of the process of division. The corollary follows by inductive use of Theorem 2 and this remark.

We are now prepared to discuss matrix functions. Let T be a measure space and $A(t)$ an nth order matrix defined for every $t \in T$ such that the n^2 scalar functions $a_{ij}(t)$ are measurable. As an immediate application of the previous theorem, there exists a measurable eigenvalue function, that is, a measurable function $\lambda(t)$ such that for every t, $\lambda(t)$ is an eigenvalue of $A(t)$.

LEMMA 3.4. *If $\lambda(t)$ is a measurable eigenvalue function then so is the dimension of its eigenspace.*

Proof. This is trivial if the rank formulation is used. The dimension is r if and only if rank $[A - \lambda I]$ is $n - r$. This can be expressed in terms of a finite number of conditions, each of which says that a certain determinant is zero or that at least one of a finite number is non-zero. Since a determinant obtained from measurable functions is measurable, the result follows.

It may be shown by calculation that vectors $[x_{i1}(t) \ldots x_{in}(t)]$ can be chosen such that for each t they form an independent basis of the eigenspace, and $x_{ij}(t)$ is measurable for all i and j.

We may now develop a canonical form for J_2 operators. For this purpose it would have sufficed to consider quadratic equations and matrices of order two, thus simplifying the proofs. However, the work was done in general, since one of our main objectives is to show that the basic ideas are valid for any n. Since unitary equivalence theory for matrices will now be used, we limit ourselves to J_2 operators. (The unitary equivalence classification of matrices of order n increases rapidly in complication as n increases. To illustrate this remark, the reader may apply the subsequent technique to J_3 operators, and discover for himself how messy the algebra becomes.)

Let A be a J_2 operator. Then we know that $R(A, A^*)$ decomposed with respect to its center gives rise to factors of type I_1 or I_2. In this decomposition A will be a scalar where the factor is of type I_1, and a block matrix D_m where D is non-normal matrix of order two where the factor is of type I_2. D can be written explicitly as

$$\begin{bmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix}.$$

A non-normal matrix of order two may have either one or two eigenvalues. If $\lambda_1(t)$ and $\lambda_2(t)$ are the two roots of the characteristic equation, then the condition of having only one eigenvalue is that $\lambda_1(t) = \lambda_2(t)$. Since $\lambda_1(t)$ and $\lambda_2(t)$ are measurable by the preceding theory, the two types each occur on measurable subsets.

Consider the set where the matrices have different eigenvalues. The larger eigenvalue $\lambda(t)$ is a measurable function of t. Also, eigenvectors $[x_1(t), x_2(t)]$ can be chosen so that both $x_1(t)$ and $x_2(t)$ are measurable. An eigenvector corresponding to D_m would be $[x_1(t), x_2(t), 0, 0, \ldots]$ which is measurable regarded as a Hilbert space function on t. Also $[-x_2(t), x_1(t), 0, 0, \ldots]$ is measurable. This gives a set of measurable functions $f_1(t), f_2(t), \ldots$ such that $A(t)$ takes on the form

$$\begin{pmatrix} \lambda(t) & a(t) & . \\ 0 & \mu(t) & . \\ . & . & . \end{pmatrix}$$

when these are used as a basis. $[A(t) f_2(t), f_1(t)]$ is measurable, and hence so is its argument $\theta(t)$. Therefore $e^{-i\theta(t)} f_{2n}(t)$ is measurable. Now if

$$f_1(t), e^{-i\theta(t)} f_2(t), \ldots$$

are used as a basis, $a(t)$ is real and positive.

In the same way functions can be chosen on the set where $A(t)$ has only one eigenvalue so that $A(t)$ will take on the form

$$\begin{pmatrix} \lambda(t) & a(t) & . \\ 0 & \mu(t) & . \\ . & . & . \end{pmatrix}$$

where $a(t)$ is real and positive if these are taken as a basis. It is well known that the matrices $D(t)$ we now have are canonical forms of the unitary equivalence classes of non-normal matrices of order two. This proves

THEOREM 3. *If A is a J_2 operator, then A can be decomposed so that for every t, $A(t)$ is either a scalar or a matrix of the form $D_m(t)$ where $D(t)$ is of the form*

$$\begin{pmatrix} \lambda & a \\ 0 & \mu \end{pmatrix}$$

where a is real and positive and $\lambda \geqslant \mu$. In particular the matrices $D(t)$ of order 2 that appear are either identical or else not even unitarily equivalent.

Consider the space $Z \cup Q$ where Z is the complex plane and Q is the set of all triples (λ, μ, a) where λ and μ are complex such that $\lambda \geqslant \mu$ and a is real and positive. We now transfer the direct integral decomposition to the space $Z \cup Q$. A projection-valued measure will be defined on $Z \cup Q$. If $B \subset Z \cup Q$ it makes sense to speak of $f^{-1}(B)$ by making the obvious identification of $\lambda \in Z$ and $(\lambda, \mu, a) \in Q$ with

$$\begin{pmatrix} \lambda & 0 & . \\ 0 & \lambda & . \\ . & . & . \end{pmatrix} \text{ and } \begin{pmatrix} \lambda & a & . \\ 0 & \mu & . \\ . & . & . \end{pmatrix}$$

respectively and then defining $f^{-1}(B) = \{t: A(t) \in B\}$. We define B to be measurable if $f^{-1}(B)$ is measurable, and then let $E(B) = E[f^{-1}(B)]$. This gives a projection-valued measure for which the Borel sets are measurable.

H decomposes with respect to the projection-valued measure on $Z \cup Q$. Also, the range of the projection-valued measure on $Z \cup Q$ is a subset of the range on the original space. We restrict ourselves for a moment to the space where $A(t)$ is not scalar. Define U_{11}, U_{12}, U_{21}, and U_{22} as in the proof of Theorem 1, for example U_{21} corresponds to the operator function which is

$$\begin{pmatrix} 0 & 0 & . & . & . \\ 1 & 0 & . & . & . \\ . & . & 0 & 0 & . \\ . & . & 1 & 0 & . \\ . & . & . & . & . \end{pmatrix}$$

for all t. Observing the original decomposition, we see that the U_{ij} commute with all the projections in the range of the measure. Therefore, a fortiori, the U_{ij} commute with all the projections in the range of the measure on $Z \cup Q$. As in the proof of Theorem 1, it is possible to choose the f's in such a way that U_{11} becomes

$$\begin{pmatrix} 1 & 0 & . & . & . \\ 0 & 0 & . & . & . \\ . & . & 1 & 0 & . \\ . & . & 0 & 0 & . \\ . & . & . & . & . \end{pmatrix}$$

on $Z \cup Q$ and similarly for U_{12}, U_{21}, and U_{22}.

Since $\|A(t)\| \in L^\infty$ in the original space, it follows that $E(Q) = 0$ outside of a Cartesian product of bounded sets.

Now subdivide Z and Q into a finite number of Borel sets $\{Z_i\}$ and $\{Q_j\}$, and choose a point $z_i \in Z_i$ and $(\lambda_j, \mu_j, a_j) \in Q_j$ for each Borel set. We define an operator function $A_\alpha(t)$ on the original space. $A_\alpha(t) = z_i$ if $A(t) \in Z_i$, and

$$A_\alpha(t) = \begin{pmatrix} \lambda_j & a_j & . \\ 0 & \mu_j & . \\ . & . & . \end{pmatrix} \text{ if } A(t) = \begin{pmatrix} \lambda(t) & a(t) & . \\ 0 & \mu(t) & . \\ . & . & . \end{pmatrix}$$

where $[\lambda(t), \mu(t), a(t)] \in Q_j$. We denote the projection on the set where $A_\alpha(t) = z_i$ by E_i and the projection on the set where

$$A_\alpha(t) = \begin{pmatrix} \lambda_j & a_j & . \\ 0 & \mu_j & . \\ . & . & . \end{pmatrix}$$

by F_j. Then A_α is simply

$$\sum_i z_i E_i + \sum_j (\lambda_j F_j U_{11} + U_j F_j U_{22} + a_j F_j U_{12})$$

in view of the definition of U_{ij}. By definition of the projection-valued measure on $Z \cup Q$, the projection on Z_i is E_i and on Q_j is F_j. It follows from this, that if A_α is decomposed with respect to $Z \cup Q$, it becomes the function which is

$$z_i \text{ on } Z_i \text{ and } \begin{pmatrix} \lambda_j & a_j & . \\ 0 & \mu_j & . \\ & . & . & . \end{pmatrix} \text{ on } Q_j.$$

We now choose a sequence of operators of the form A_α which approach A uniformly. (Note that the property of uniform convergence is preserved under direct integrals.) By using this sequence, we easily see the important fact that if A is decomposed with respect to $Z \cup Q$ it becomes the function which is

$$\begin{pmatrix} \lambda & 0 & . \\ 0 & \lambda & . \\ & . & . & . \end{pmatrix} \text{ at } \lambda \text{ and } \begin{pmatrix} \lambda & a & . \\ 0 & \mu & . \\ & . & . & . \end{pmatrix} \text{ at } (\lambda, \mu, a).$$

THEOREM 4 (the spectral theorem). *If A is a J_2 operator, then there exists a direct integral decomposition onto $Z \cup Q$ such that A is decomposable and such that*

$$A(t) = \begin{pmatrix} \lambda & 0 & . \\ 0 & \lambda & . \\ & . & . & . \end{pmatrix} at \ \lambda \ and \ \begin{pmatrix} \lambda & a & . \\ 0 & \mu & . \\ & . & . & . \end{pmatrix} at \ (\lambda, \mu, a).$$

It is natural to call this "the spectral theorem" and even to use the symbolic notation $A = \int \lambda \, dE(\lambda)$ because, at least from the point of view of direct integral theory, it generalizes the usual theorem for normal operators.

4. Uniqueness of spectral representations. In this section we show that the measure obtained on $Z \cup Q$ is unique when A is given. To simplify the notation, the matrices $A(t)$ will be written as if they have only one block. This is legitimate since any ring of the form

$$\begin{pmatrix} R & 0 & . \\ 0 & R & . \\ & . & . & . \end{pmatrix}$$

is isomorphic and isometric to R.

Since we are given that the identity function λ corresponds to A, we know by direct integral theory that $f(\lambda, \lambda^*)$ corresponds to $f(A, A^*)$ where f is a polynomial function. The general idea will be to approximate characteristic functions by such polynomials, not uniformly, but closely enough to ensure that the projections on the sets depend only on A. For normal operators, the Stone-Weierstrass approximation theorem can be immediately applied and the

proof is rather trivial. As Kaplansky **(3)** has shown, the theorem can be generalized to the case where the range is in a B^* algebra, but a certain amount of caution must be used in applying the theorem.

For example, separation of points is not enough. In fact, consider a compact set of matrices of order two in the usual topology. We remark that if the set consists of two distinct matrices which are unitarily equivalent, then the B^* algebra generated by the identity function is not dense in the space of all continuous functions even though the algebra separates points. This is because any scalar function in the algebra necessarily has the same value at unitarily equivalent matrices. Kaplansky's theorem asserts that if any two points a and b can be separated in the sense that there exists an f such that $f(a) = 1$ and $f(b) = 0$, then the algebra is dense in the set of all continuous scalar functions.

THEOREM 5. *If x and y are any two distinct points in $Z \cup Q$, then there exists a polynomial function f such that $f(x) = 1$ and $f(y) = 0$. (* is regarded as a polynomial operation.)*

Proof. We regard x and y as matrices in the natural manner and consider the matrix

$$\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}.$$

The structure of the space $Z \cup Q$ ensures that x and y are not unitarily equivalent. Suppose there exists a polynomial function f such that

$$f\left[\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}\right] = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then it follows that $f(x) = 1$ and $f(y) = 0$. Thus the problem has been reduced to the study of the algebra of polynomials on a matrix of the type

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

where A and B are not unitarily equivalent, and each of them together with its adjoint generates the full matrix ring of which it is a member. The latter part of the statement is true by definition of $Z \cup Q$. (For example, Q does not contain points such as $(\lambda, \mu, 0)$.)

It is known that there exists a polynomial function f such that

$$f\left[\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right] = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix},$$

where U is any matrix in the full matrix ring of which A is a member and similarly for V. This was proved independently in the thesis **(2)** by making use of the classification theorem for finite dimensional rings of operators as found in **(5)**. The result follows by taking $U = 1$ and $V = 0$.

Although the question of separation of points has been answered, other difficulties arise. Q is not compact in its usual topology. The addition of points $(\lambda, \mu, 0)$ will make Q compact (if we assume that $|\lambda|, |\mu|, a \leqslant ||A||$) but will destroy the property of separation of points. For example, it is clear that the points $(\lambda, \mu_1, 0)$ and $(\lambda, \mu_2, 0)$ with $\mu_1 \neq \mu_2$ could not be separated, for if

$$f\begin{pmatrix} \lambda & 0 \\ 0 & \mu_1 \end{pmatrix} = 1 \text{ and } f\begin{pmatrix} \lambda & 0 \\ 0 & \mu_2 \end{pmatrix} = 0,$$

then $f(\lambda) = 1$ and $f(\lambda) = 0$ which is impossible. Accordingly we regard Q as a locally compact space and search for "enough" functions vanishing at infinity.

Definition. Let X be an arbitrary locally compact space. A family of compact subsets k_α is said to form an *inverse base* for compact sets if every compact subset of X is contained in at least one k_α.

Example. In the Euclidean plane, a countable inverse base for compact sets can be obtained by taking circles of integral radius and center at the origin.

The concept of an inverse base is useful, because many properties necessarily hold for all compact sets if they hold for sets of a base (which are often easier to handle than arbitrary compact sets). For example, the condition for a normal family of functions of a complex variable can be expressed in terms of sets of an inverse base.

We restrict ourselves to the subset of Q where $|\lambda| \leqslant ||A||$, $|\mu| \leqslant ||A||$, and $a \leqslant ||A||$. In any decomposition of A it is clear that the measure is concentrated on this subset. No confusion should arise if Q is used to denote this subset.

LEMMA 4.1. *The sets Q_α, where α runs through all positive real numbers and Q_α is the set of all $(\lambda, \mu, a) \in Q$ such that $a \geqslant \alpha$, form an inverse base for compact sets of Q.*

Proof. Clearly Q_α is compact for all α. Let k be a compact subset of Q. The mapping $(\lambda, \mu, a) \to a$ is continuous, and therefore has a minimum α when restricted to k. Obviously $k \subset Q_\alpha$.

COROLLARY. *A function on Q vanishes at infinity if and only if it approaches zero, as a approaches zero, uniformly with respect to λ and μ.*

LEMMA 4.2. *All polynomials of the form $fg - gf$ vanish at infinity.*

Proof.

$$(fg - gf)\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} = \begin{bmatrix} (fg - gf)\lambda & 0 \\ 0 & (fg - gf)\mu \end{bmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Also $fg - gf$ is continuous on $Q \cup (\lambda, \mu, 0)$, and is hence uniformly continuous. The rest follows by the corollary to Lemma 4.1.

Now let x and y be two distinct points in Q. By the strong form of Theorem 5 (see the end of the proof), there exists a polynomial f such that

$$f(x) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } f(y) = 0.$$

Consider $(ff^* - f^*f)^2$. By Lemma 4.2, this vanishes at infinity since algebraic combinations of functions vanishing at infinity still vanish at infinity. The function is clearly zero at y and computation shows that it is one at x. This proves

LEMMA 4.3. *If x and y are any two distinct points in Q, there exists a polynomial function f vanishing at infinity such that $f(x) = 1$ and $f(y) = 0$.*

Notice that not all polynomial functions on Q vanish at infinity, but we have shown that those that do vanish, separate points. By Kaplansky's theorem (3) the set of polynomial functions that vanish at infinity is dense in the set of all scalar functions vanishing at infinity.

Next we approximate characteristic functions by continuous functions. It is clear that the Borel sets in Q are generated by the compact sets. Let k be a compact set. Since Q is a metric space, k' is a countable union of closed sets $F_1 \subset F_2 \subset F_3 \subset \dots$. By applying Urysohn's lemma to the F_i in succession we obtain continuous functions which are uniformly bounded (in fact by 1) and which tend pointwise to the characteristic function on k. By means of these results we can now solve the uniqueness problem stated at the beginning of the section.

We know that $f(\lambda, \lambda^*)$ corresponds to $f(A, A^*)$ for all polynomial functions f. Since all continuous scalar functions vanishing at infinity can be uniformly approximated by polynomials, they correspond to well defined limits of sequences of operators each of which is of the form $f(A, A^*)$. A pointwise limit of a uniformly bounded sequence of functions corresponds to the strong limit of the sequence of operators corresponding to the functions by direct integral theory. This means that the projections on compact sets are unique. This implies uniqueness of the projection-valued measure restricted to Q. In particular, the function which is 1 on Q corresponds to a well-defined projection.

So far, all functions considered were 0 on Z. Now by subtraction, the function which is 1 on Z and 0 on Q corresponds to a unique projection. (Of course, the function which is 1 everywhere corresponds to 1 and therefore to something unique.) Multiplying by the identity function λ which corresponds to A, we see that the function which is the identity on Z (λ at λ) and 0 on Q corresponds to a unique operator. It is now easy to verify that the projection-valued measure is unique on Z. (The proof is essentially the same as the proof for uniqueness on Q except that the details are much simpler for Z.)

Before stating the fundamental theorem it is convenient to make a definition.

Definition. If E is a projection-valued measure on $Z \cup Q$,

$$A = \int \lambda \, dE(\lambda)$$

means that if A is decomposed with respect to the measure, it becomes an operator function of the form

$$\begin{pmatrix} \lambda & a & . & . & . \\ 0 & \mu & . & . & . \\ . & . & \lambda & a & . \\ . & . & 0 & \mu & . \\ . & . & . & . & . \end{pmatrix} \text{ at } (\lambda, \mu, a) \text{ and } \begin{pmatrix} \lambda & 0 & . \\ 0 & \lambda & . \\ . & . & . \end{pmatrix} \text{ at } \lambda.$$

THEOREM 6. *There is a one-one correspondence between J_2 operators and projection-valued measures on $Z \cup Q$ concentrated on a compact set such that if A corresponds to E,*

$$A = \int \lambda \, dE(\lambda).$$

As a corollary we have the theorem of unitary equivalence.

THEOREM 7. *There is a one-one correspondence between unitary equivalence classes of J_2 operators and collections of null sets which include the complement of some compact set together with a measurable multiplicity function where the range consists of the positive integers and the symbol $\aleph_0$, and the range is even valued on Q; where the correspondence is obtained by Theorem 6 together with multiplicity theory.*

Remark 1. If we have a decomposition of A in the manner described in Theorem 6, then the decomposition is necessarily with respect to the center of $R(A, A^*)$ i.e. the center of $R(A_1 A^*)$ is precisely the set of all operators which decompose into scalar functions.

Remark 2. Not only does A determine the E's uniquely but even the U_{ij}'s are uniquely determined. (The U_{ij}'s are used in the same sense as before with, for example, U_{11} corresponding to the function which is

$$0 \text{ on } Z \text{ and } \begin{pmatrix} 1 & 0 & . \\ 0 & 0 & . \\ . & . & . \end{pmatrix} \text{ on } Q.$$

To verify the second remark it is necessary to make use of another theorem by Kaplansky **(3)** which also considers functions with range in a B^* algebra. The theorem asserts that if any two points x and y can be separated in the sense that for any two elements G and H in B^*, there exists an f such that $f(x) = G$ and $f(y) = H$, then the algebra is dense in the set of all continuous functions with range in the B^* algebra.

Thus a J_2 operator has a rich supply of projections associated with it, "vertical" (such as U_{11}) as well as "horizontal."

5. Applications. Brown **(1)** obtained a unitary equivalence theory for a class of "binormal" operators. He approached the problem in a purely algebraic manner completely avoiding direct integral theory. We now state without

proof various relations between his approach and ours. More detail can be found in **(2)**.

A is binormal $\rightleftarrows A$ is a J_2 operator.

The normal kernel of A (according to Brown) is the projection on Z.

$A^2 = 0 \rightleftarrows$ measure on $Z \cup Q$ associated with A is concentrated on 0, and $(0, 0, a)$.

$A^2 = A \rightleftarrows$ measure on $Z \cup Q$ associated with A is concentrated on (0), (1), and $(1, 0, a)$.

The last two results enable us to find a unitary equivalence theory for nilpotent (of order two) and idempotent operators.

It is useful to have other forms of the spectral theorem which do not involve direct integral theory or refer to the specific space $Z \cup Q$.

If A is any J_2 operator, then there exists a sequence $\{A_n\}$ of operators approaching A uniformly such that for all n, A_n has the property that the Hilbert space can be split into a finite number of orthogonal spaces on each of which it is of the form

$$\begin{pmatrix} \lambda & 0 & . \\ 0 & \lambda & . \\ . & . & . \end{pmatrix} \text{ or } \begin{pmatrix} \lambda & a & 0 & 0 & . \\ 0 & \mu & 0 & 0 & . \\ 0 & 0 & \lambda & a & . \\ 0 & 0 & 0 & \mu & . \\ . & . & . & . & . \end{pmatrix}$$

Note that for J_1 operators this corresponds to the well-known fact that A can be uniformly approximated by operators of the form

$$\sum_{i=1}^{n} \lambda_i E_i.$$

6. Generalizations.

The following generalizations are possible.

(1) We may define unbounded J_2 operators. The results will still be valid except for the fact that the measure is not concentrated on a compact set.

(2) The main theorems can be extended to the case where $n > 2$. The chief difficulty is that instead of the space $Z \cup Q$, we shall have $Z \cup Q \cup Q_3 \cup Q_4 \ldots \cup Q_n$ where the spaces increase rapidly in complication. (Even Q_3 is complicated!) However, the basic ideas still go through.

References

1. A. Brown, *On binormal operators*, Amer. J. Math., *76* (1954), 414–434.
2. H. Gonshor, *Spectral theorem for a class of non-normal operators*, Ph.D. thesis, Harvard University, June 1953.
3. I. Kaplansky, *The structure of certain operator algebras*, Amer. Math. Soc. Trans., *70* (1951), 219–233.
4. B. E. Mitchell, *Unitary transformations*, Can. J. Math., *6* (1954), 69–72.
5. J. von Neumann, *On rings of operators, reduction theory*, Ann. Math., *50* (1949), 401–485.
6. J. von Neumann and F. J. Murray, *On rings of operators* IV, Ann. Math., *44* (1943), 772–773.

Pennsylvania State University