SPECTRAL THEORY OF NORMAL OPERATORS ON HILBERT SPACE

A THESIS

Presented to

The Faculty of the Graduate Division

by

Monte Alan Franklin

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Mathematics

Georgia Institute of Technology

June, 1968

In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

3/17/65
b

SPECTRAL THEORY OF NORMAL OPERATORS

ON HILBERT SPACE

Approved:

Chairman

Date approved by Chairman: May 27/68

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to his teacher and advisor, Dr. Eric R. Immel, for his guidance and valuable suggestions, and for much inspiration. The author also wishes to express his thanks to Dr. James M. Osborn and Dr. Harold A. Gersch for reading the manuscript and making numerous helpful suggestions. It is a pleasure to acknowledge the assistance and support the author received from The MITRE Corporation during the preparation of the manuscript. He is especially grateful to Elizabeth Couilliard for typing the manuscript.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1. A Historical Note

The spectral theorem asserts, in essence, that a bounded normal operator on Hilbert space can be approximated in the operator norm by linear combinations of projection operators (a "spectral resolution" or "decomposition" of the bounded normal operator). Since the inauguration of modern trends in spectral theory by J. von Newmann, M. H. Stone and others, proofs of the spectral theorem have been given in various settings for Hermitian, unitary, and normal operators. While the spectral theorem for Hermitian operators is comparatively easy to prove, the same result for normal operators is much more difficult.

The "algebraization" of the spectral theory was initiated by J. von Newmann [20] (1936), H. Freudenthal [6] (1936) and S. Steen [26] (1936). M. Nagumo [18] (1936) contributed the notion of Banach algebras to analysis. The tendency to emphasize the algebraic aspects of the spectral theory was continued in the work of S. Kakutani [12] (1939), F. Riesz [23] (1940), M. H. Stone [27] (1940, 1941) and B. Vulich [33] (1940). The ideal theory of normed rings due to I. M. Gelfand [7] (1941) proved to be vitally important in following developments. K. Yosida and T. Nakayama [32] (1942) applied the Gelfand theory in proving an "abstract spectral theorem" for elements of certain semi-ordered rings, and obtained the spectral theorem for bounded Hermitian operators on Hilbert space as a special case. A proof of the spectral theorem for bounded normal operators (independent of the Hermitian case) was given by K. Yosida [29] (1943), where the ideal theory of normed rings was applied directly to a commutative ring of normal operators.

Note. In much of the literature the terminology "ring" is used in place of "algebra," for the algebraic structure with a multiplication defined. In this thesis the later terminology will be adopted. In this connection, we shall speak of the "Gelfand theory of Banach algebras," rather than "the ideal theory of normed rings."

2. Summary

The aim of this thesis is to give the main prerequisites for a study of abstract spectral theory from the algebraic viewpoint and a proof of the spectral theorem for normal operators in this setting. It is intended that the exposition should be readable on the basis of a good general background in real analysis. Throughout this thesis references are provided for the more basic, well-known results, in lieu of proofs.

Chapter II consists of the more fundamental definitions and facts concerning Hilbert spaces, operators on a Hilbert space, and Hermitian operators, in particular. The main results of this chapter, which are needed later on, are the supremum formula for the norm of a Hermitian operator T,

$$\|T\| = \sup\{|(Tx,x)| : \|x\| \leq 1\} ;$$

the partial ordering $\leq$ on the class of Hermitian operators; the Hermitian decomposition of an operator; and the fact that the operators $I + T^*T$, $I + TT^*$ are invertible, where T is Hermitian.

Chapter III gives pertinent results from the Gelfand theory of Banach algebras. Section 1 deals with the general properties of complex commutative algebras with unit, maximal ideals, homomorphisms, and quotient algebras modulo a maximal ideal. In order to fully develop the Gelfand theory of Banach algebras, we shall make use of a generalized concept of analytic function. Section 2 indicates how the theory of ordinary analytic functions is extended to cover "abstract analytic functions" which have values in Banach space. The abstract versions of the Liouville theorem, the Cauchy integral theorem, the Cauchy integral formula, and the Taylor series expansion of an analytic function are given. The proofs of the theorems cited above in their abstract versions rest on a consequence of the Hahn-Banach theorem. The principal results of Section 3 are the Gelfand-Mazur theorem, the characterization of a commutative Banach algebra A with unit by the Gelfand representation, and the Spectral Radius Formula,

$$\rho(x) = \lim_{n \to \infty} \|x^n\|^{1/n} \quad (x \in A) \quad .$$

The Gelfand topology on a maximal ideal space is discussed in Section 4. In this concluding section of Chapter III we also prove the Alaoglu theorem, and the fact that the structure space of a commutative Banach algebra with unit (the maximal ideal space with the Gelfand topology) is a compact Hausdorff space. The Tychonoff theorem on product topologies is used in the proof of the Alaoglu theorem.

In Chapter IV the Gelfand theory of Banach algebras is applied to a certain commutative Banach algebra of normal operators on Hilbert space. This particular Banach algebra, called "the Banach algebra B generated by a normal operator," is constructed in Section 1. The fundamental properties of the Gelfand representation of the Banach algebra B are proved in Section 2. An important conclusion in this development, proved by using a consequence of the Stone-Weierstrass theorem, states that the space of continuous functions which represents the Banach algebra B by the Gelfand representation is actually equal to the totality of continuous functions on the structure space of B. We relate convergence in norm (strong limit) of a monotone sequence of positive operators in B to convergence (uniform limit) of a sequence of continuous functions on the structure space of B in Section 3, by using two theorems due to Baire. The spectral resolution of a normal operator is constructed in the concluding section.

3. Notation

Within this thesis, items (definitions, lemmas, theorems, corollaries, and remarks) are numbered consecutively within a section. Chapters are referred to by roman numerals, sections and items are referred to by arabic numerals. Thus, "Theorem III.3.10" refers to Theorem 10 of Section 3(the tenth theorem of Section 3) in Chapter III. If no chapter number appears in a reference then the item referenced is within the same chapter. Similarly, if no section number appears in a reference, then the item referenced is within the same section. Thus, "Remark 1" refers to the first remark in the section in which the reference "Remark 1" appears. Equations, inequalities and other conditions are also numbered consecutively within a section, with no connection to the numbering of the items mentioned previously. These conditions will be referenced by a number in parenthesis. For example, "by III.3(14)" refers to the condition numbered (14) in Section 3 of Chapter III.

Bibliographical references are cited by the name of the author followed by the number of the reference in the alphabetized BIBLIOGRAPHY in square brackets.

CHAPTER II

BACKGROUND IN HILBERT SPACE THEORY

1. Hilbert Spaces

Definitions. A linear space H over the complex field C (the scalar field) is an inner product space (or pre-Hilbert space) if there is a mapping which makes correspond to each pair $x,y \in H$ a complex number (x,y), called the inner product of x and y, which satisfies the following conditions for all $x,y,z \in H$ and all $\lambda \in C$:

$$(x,x) \geq 0, \text{ and } (x,x) = 0 \text{ if and only if } x = 0 \tag{1}$$

$$(x + y,z) = (x,z) + (y,z) \tag{2}$$

$$(\lambda x,y) = \lambda(x,y) \tag{3}$$

$$\overline{(x,y)} = (y,x), \text{ where } \overline{(x,y)} \text{ is the} \tag{4}$$

complex conjugate of (x,y).

It is easily verified that $(x,y + z) = (x,y) + (x,z)$ and $(x,\lambda y) = \bar{\lambda}(x,y)$.

If H is an inner product space, a norm is introduced in H by defining $\|x\| = (x,x)^{\frac{1}{2}}$. It follows from properties of the inner product that, for all $x,y \in H$ and all scalars λ,

$$\|x\| \geq 0, \text{ and } \|x\| = 0 \text{ if and only if } x = 0 \tag{5}$$

$$\|\lambda x\| = |\lambda| \cdot \|x\| \tag{6}$$

$$|(x,y)| \leq \|x\| \cdot \|y\| \quad \text{(Schwarz inequality)} \tag{7}$$

$$\|x + y\| \leq \|x\| + \|y\| \quad \text{(Minkowski inequality)} . \tag{8}$$

The Schwarz inequality is proved in Rudin [25], for example. This norm, in turn, induces a metric d on H defined by $d(x,y) = \|x-y\|$, called the distance between x and y. An open ball in H with center x and radius $r > 0$ is denoted by $S(x,r)$. Thus

$$S(x,r) = \{y : y \in H, \|x - y\| < r\} . \tag{9}$$

The corresponding metric topology in H is obtained by specifying the class of open sets as follows. A set $E \subset H$ is an open set in H if to each point $x \in E$ there corresponds an open sphere $S(x,r)$ (with $r > 0$ depending on x) such that $S(x,r) \subset E$. A sequence $\{x_n\}$ in H is a Cauchy sequence (in the metric space H) if, for each $\epsilon > 0$, there exists a positive integer N such that $\|x_m - x_n\| < \epsilon$ whenever $m,n > N$. An inner product space which is complete as a metric space (with the metric induced by the inner product, as outlined above) is called a Hilbert space. An inner product space H is complete if, whenever $\{x_n\}$ is a Cauchy sequence in H. there exists an element $x \in H$ such that

$$\lim_{n \to \infty} \|x_n - x\| = 0 \quad . \tag{10}$$

We now mention some examples of Hilbert spaces. Proofs are omitted in this section.

<u>Examples</u>. (a). Suppose (X,M,μ) is a measure space, with μ a positive measure on a σ-algebra M of subsets of X. Let $S^2(X,\mu)$ be the class of all complex-valued measurable functions f on X such that $|f|^2$ is <u>integrable</u>. This class is a complex linear space. Two functions f, g, $\in S^2(X,\mu)$ are called <u>μ-equivalent</u> if $f(x) = g(x)$ except on a set of μ-measure zero. By this equivalence relation, $S^2(X,\mu)$ is split into disjoint equivalence classes. Let $\hat{f}$ be the equivalence class containing the function f, that is,

$$\hat{f} = \{g : g \in S^2(X,\mu);\ f,\ g \text{ are } \mu\text{-equivalent}\} \quad .$$

The space of equivalence classes (which is a complex linear space if we define $\hat{f} + \hat{g} = \widehat{f + g}$, $\lambda\hat{f} = \widehat{\lambda f}$), is called $\mathcal{L}^2(X,\mu)$ when it is understood as equipped with the norm

$$\|\hat{f}\| = \left[\int_X |f|^2 \, d\mu\right]^{\frac{1}{2}} \quad . \tag{11}$$

If $\hat{f}, \hat{g} \in \mathcal{L}^2(X,\mu)$, let

$$(\hat{f},\hat{g}) = \int_X f \, \bar{g} \, d\mu \quad . \tag{12}$$

In this way, an inner product is specified, and it can be proved that $\mathcal{L}^2(X,\mu)$, with this inner product, is a Hilbert space. For the detailed arguments necessary in this example, see for instance Rudin [25], pp. 66-67. In practice, and in much of the literature, less precise notation is employed, and no explicit mention is made of equivalence classes. With the less precise notation, one simply refers to the <u>function space</u> $S(X,\mu)$ as the Hilbert space, and agrees to "identify" μ-equivalent functions.

(b). Let ℓ^2 be the space of all sequences $\{a_n\}$ of complex numbers such that

$$\sum_{n=1}^{\infty} |a_n|^2 < \infty \quad .$$

Define

$$(a,b) = \sum_{n=1}^{\infty} a_n \overline{b}_n \quad \text{if } a = \{a_n\},\ b = \{b_n\} \quad .$$

Then ℓ^2 can be shown to be a Hilbert space.

2. Operators on a Hilbert Space

Throughout this thesis, the term <u>operator on H</u> will refer to a <u>linear transformation</u> T mapping H into H which is <u>bounded</u> in the sense that

$$\sup\{\|Tx\| : x \in H, \|x\| \leq 1\} \leq \infty \quad . \qquad (1)$$

The set of all operators on H will be denoted by A(H). For all $S,T \in A(H)$ and all scalars λ, define λT, $S + T$ (operator sum), and ST (operator product) by

$$(\lambda T)x = \lambda(Tx) , \tag{2}$$

$$(S + T)x = Sx + Tx , \tag{3}$$

$$(ST)x = S(Tx) , \tag{4}$$

for every $x \in H$. It is easily verified that λT, $S + T$, and ST are elements of A(H), and that A(H) is a complex linear space under the scalar multiplication and addition defined. Since, in addition, a multiplication is defined under which

$$S(T_1 + T_2) = ST_1 + ST_2 \quad \text{and} \tag{5}$$

$$(S_1 + S_2)T = S_1T + S_2T , \tag{6}$$

A(H) is called an <u>algebra</u> <u>of</u> <u>operators</u>. A(H) contains the <u>identity</u> <u>operator</u> I defined by $Ix = x$ for all $x \in H$. The <u>zero</u> <u>operator</u> O of the linear space A(H) is defined by $Ox = 0$ for all $x \in H$. If n is a positive integer, $T^n = T \cdot \cdot \cdot T$, the product of T which itself n times. It is convenient to set $T^o = I$ for each $T \in A(H)$. If

$$p(z) = \sum_{j=0}^{n} \alpha_j z^j ,$$

is a polynomial in z with complex coefficients, it is easily verified that p(T), defined by

$$p(T) = \sum_{j=0}^{n} \alpha_j T^j ,$$

is contained in A(H) if $T \in A(H)$.

Definition 1. The norm of an operator T, written $\|T\|$, is the nonnegative real number

$$\|T\| = \sup\{\|Tx\| : x \in H, \|x\| \le 1\} .$$

By (1), $\|T\|$ is finite for each $T \in A(H)$. Since a linear transformation is bounded if and only if it is continuous, every element of A(H) is a continuous function on H. The norm of an operator T is also given by the formulas (see Rudin [25], p. 96; Berberian [2], pp. 94-96)

$$\|T\| = \sup\{ \frac{\|Tx\|}{\|x\|} : x \in H, x \neq 0\}, \text{ if } H \neq \{0\} \tag{7}$$

$$\|T\| = \sup\{\|Tx\| : x \in H, \|x\| = 1\}, \text{ if } H \neq \{0\} \tag{8}$$

$$\|T\| = \inf\{\gamma : \|Tx\| \le \gamma\|x\| \text{ for all } x \in H\} . \tag{9}$$

Theorem 1. Let S and T be operators on H, and let α be a complex number. Then

$$\|\alpha T\| = |\alpha| \cdot \|T\| \tag{10}$$

$$\|S + T\| \leq \|S\| + \|T\| \tag{11}$$

$$\|ST\| \leq \|S\| \cdot \|T\| \quad . \tag{12}$$

This theorem is proved in Halmos [9], p. 35.

Properties (11) and (12), and the fact that $\|T\| = 0$ if and only if $T = 0$, show that $\|\cdot\|$ is a norm on the space A(H). The linear space A(H) is complete in the norm defined; that is, A(H) is a Banach space. (Refer, for example, to Taylor [28], p. 163, or to Hewitt and Stromberg [10], p. 211.)

Definition 2. An operator $T \in A(H)$ is said to be invertible if there exists an operator $S \in A(H)$ such that $ST = TS = I$. If such an operator S exists for $T \in A(H)$, then S is unique, and is called the inverse operator of T, and denoted by T^{-1}. An operator which is not invertible is said to be singular.

Note. If $ST = TS = I$ and $S_1T = TS_1 = I$, then $S = IS = (S_1T)S = S_1(TS) = S_1I = S_1$.

Definition 3. Let T be an operator on H. The set of all complex numbers λ such that $T - \lambda I$ is singular is called the spectrum of the operator T, and is denoted by $\sigma(T)$. The complement of $\sigma(T)$ in the complex plane C is called the resolvent set of T, denoted by r(T).

3. Hermitian Operators on a Hilbert Space

Definition 1. Let $T \in A(H)$. If there exists an operator T^* such that $(y,Tx) = (T^*y,x)$ for all $x,y \in H$, then T^* is called the adjoint operator of T, or the Hilbert space adjoint of T.

The existence of a unique adjoint operator T^* for every operator $T \in A(H)$ can be proved as a consequence of the Riesz-Fréchet representation theorem for bounded linear functionals on a Hilbert space. (See, for example, Halmos [9], pp. 17, 38.)

Theorem 1. Let $S,T \in A(H)$ and $\lambda \in C$. Then

$$T^{**} = T \quad (1)$$

$$(\lambda T)^* = \bar{\lambda} \cdot T^* \quad (2)$$

$$(S + T)^* = S^* + T^* \quad (3)$$

$$(ST)^* = T^* S^* \quad . \quad (4)$$

If S is invertible, then S^* is invertible, and $(S^*)^{-1} = (S^{-1})^*$. (5)

A proof of this theorem is given in Halmos [9], p. 39.

Definition 2. An operator T on H is said to be Hermitian (or self-adjoint) if $T^* = T$, and normal if $TT^* = T^*T$.

Remark 1. Note that a real scalar multiple of a Hermitian operator is Hermitian, and the sum of two Hermitian operators is Hermitian.

The class of all Hermitian operators on H forms a real normed linear space which will be denoted B(H). If T is a Hermitian operator on H, then for any element $x \in H$

$$\overline{(Tx,x)} = (x,Tx) = (x,T^*x) = (Tx,x) \quad . \tag{6}$$

Thus, if $T \in A(H)$ is Hermitian, (Tx,x) is <u>real</u> for every $x \in H$. The converse is also true; that is, if (Tx,x) is <u>real</u> for every $x \in H$, then T is necessarily Hermitian. This is proved as follows.

Suppose that $T \in A(H)$, and (Tx,x) is real for every $x \in H$. Then $(Tx,x) = (x,Tx) = (T^*x,x)$ for every $x \in H$. Hence $((T - T^*)x,x) = 0$ for every $x \in H$. Let $S = T - T^*$. It follows from the standard <u>polarization identity</u>

$$4(Sx,y) = (S(x + y), x + y) - (S(x - y), x - y)$$

$$+ i(S(x + iy), x + iy) - i(S(x - iy), x - iy)$$

that $(Sx,y) = 0$ for all $x,y \in H$. Thus $(Sx,Sx) = \|Sx\|^2 = 0$ for every $x \in H$, and $Sx = 0$ for every $x \in H$. Hence $S = \Theta = T - T^*$, and T is Hermitian.

The above remarks demonstrate the usefulness of the set of numbers $\{(Tx,x) : x \in H\}$, where T is a Hermitian operator on H. If $(Tx,x) \geq 0$ for every $x \in H$, we shall call T (necessarily Hermitian) a <u>positive</u> operator, and the notation $0 \leq T$ (or $T \geq 0$) will be used to denote this property of T. If S, T are Hermitian operators on H, then $S \leq T$

(or $T \geq S$) will be used to mean that $T - S$ is a positive operator.

As we shall see in the following theorem, the supremum of all the values taken by $|(Tx,x)|$ over the unit ball $\{x : \|x\| \leq 1\}$ in H, where $T \in A(H)$ is Hermitian, is equal to the norm of the Hermitian operator T.

Theorem 2. If T is a Hermitian operator on H, then

$$\|T\| = \sup\{|(Tx,x)| : x \in H, \|x\| \leq 1\} \quad . \tag{7}$$

Proof. Put $\gamma = \sup\{|(Tx,x)| : x \in H, \|x\| \leq 1\}$. By the Schwarz inequality, if $\|x\| \leq 1$ then

$$|(Tx,x)| \leq \|Tx\| \cdot \|x\| \leq \|T\| \cdot \|x\|^2 \leq \|T\| \quad ,$$

and hence $\gamma \leq \|T\|$. To show that $\|T\| \leq \gamma$, note that for each real number β and all $x,y \in H$, the inequality

$$|(T(x \pm \beta y), x \pm \beta y)| \leq \gamma \|x \pm \beta y\|^2 \tag{8}$$

is valid. Also, since T is Hermitian,

$$(Ty,x) = (y,Tx) = \overline{(Tx,y)} \quad . \tag{9}$$

Using (8) and (9) we find that

$$|(T(x + \beta y), x + \beta y) - (T(x - \beta y), x - \beta y)| = |2\beta(Tx,y) + 2\beta\overline{(Tx,y)}|$$

$$= \left| 4\beta \ \mathrm{Re}(Tx,y) \right|$$

and $$\left| (T(x + \beta y),\ x + \beta y) - (T(x - \beta y),\ x - \beta y) \right|$$

$$\leq \gamma(\|x + \beta y\|^2 + \|x - \beta y\|^2) = \gamma((x + \beta y,\ x + \beta y) - (x - \beta y,\ x - \beta y))$$

$$= 2\gamma(\|x\|^2 + \beta^2\|y\|^2) \quad .$$

Hence the inequality

$$\left| 2\beta \ \mathrm{Re}(Tx,y) \right| \leq \gamma(\|x\|^2 + \beta^2\|y\|^2) \qquad (10).$$

is valid for every real β and every $x,y \in H$. If $y \neq 0$, and β is replaced by $\|x\|/\|y\|$ in (10), the inequality

$$\left| \mathrm{Re}(Tx,y) \right| \leq \gamma\|x\|\cdot\|y\| \qquad (11)$$

is obtained. Note that (11) holds trivially in case $x = 0$. This inequality also holds for $y = 0$, and thus for all $x,y \in H$. If the complex number (Tx,y) is written in polar form $|(Tx,y)|e^{i\alpha}$ (where α is real), then

$$e^{-i\alpha}(Tx,y) = |(Tx,y)| \geq 0 \quad ,$$

and hence

$$\mathrm{Re}\left\{e^{-i\alpha}(Tx,y)\right\} = |(Tx,y)| \quad . \tag{12}$$

If y is replaced by $e^{i\alpha}y$ in (11), we obtain

$$\mathrm{Re}\left\{(Tx,\ e^{i\alpha}y)\right\} \leq \gamma\|x\|\cdot\|y\| \quad ,$$

since $|e^{i\alpha}| = 1$. By a property of the inner product and the above inequality, we obtain

$$\mathrm{Re}\left\{e^{-i\alpha}(Tx,y)\right\} \leq \gamma\|x\|\cdot\|y\| \quad .$$

In view of (12), it follows that

$$|(Tx,y)| \leq \gamma\|x\|\cdot\|y\| \quad .$$

In particular,

$$\|Tx\|^2 = (Tx,x) \leq \gamma\|x\|\cdot\|Tx\| \quad ,$$

and consequently $\|Tx\| \leq \gamma\|x\|$, since this inequality holds trivially if $\|Tx\| = 0$, while if $\|Tx\| > 0$ it is an immediate consequence of the preceding inequality. Thus the set $\{\|Tx\| : x \in H,\ \|x\| \leq 1\}$ is bounded above by γ, and (by definition of $\|T\|$)

$$\|T\| \leq \gamma \quad .$$

Therefore $\|T\| = \gamma$, since it has already been proved that $\gamma \leq \|T\|$.

Remark 2. The hypothesis of Theorem 2 can be relaxed considerably. It has been proved (see Bernau and Smithies [4])that for any normal operator T on H,

$$\|T\| = \sup\{|(Tx,x)| : \|x\| \leq 1\} .$$

Theorem 3. The set B(H) is partially ordered by the relation $\leq$.

Proof. Let S, T and U be elements of B(H). Then $S \leq S$, since $(Sx,x) - (Sx,x) = 0 \geq 0$ for every $x \in H$ implies that $S - S \geq 0$. Next, suppose that $S \leq T$ and $T \leq U$. It must be shown that $S \leq U$. Note that $(Ux,x) \geq (Tx,x)$ and $(Tx,x) \geq (Sx,x)$ for each $x \in H$, and hence $S \leq U$. Finally, suppose that $S \leq T$ and $T \leq S$. It must be shown that $S = T$. Note that $(Sx,x) \leq (Tx,x)$ and $(Tx,x) \leq (Sx,x)$ for each $x \in H$ implies that $(Sx,x) = (Tx,x)$ for each $x \in H$. That is,

$$(Tx - Sx,x) = 0 = ((T-S)x,x)$$

for every $x \in H$. By Theorem II.3.2 this implies that $\|T - S\| = 0$, which implies that

$$T - S = 0 .$$

Remark 3. If $T \in A(H)$, then T can be decomposed into "real" and "imaginary" parts as follows. Define

$$R = \frac{1}{2}(T + T^*) \text{ and } S = \frac{1}{2i}(T - T^*) \quad . \tag{13}$$

By Theorem 1,

$$R^* = \tfrac{1}{2}(T + T^*)^* = \tfrac{1}{2}(T^* + T^{**}) = \tfrac{1}{2}(T^* + T) = R.$$

Similarly $S^* = S$. Thus, R and S are Hermitian operators. If $T = R_1 + iS_1$, where R_1 and S_1 are Hermitian operators, it follows from Theorem 1 that $T^* = R_1^* - iS_1^* = R_1 - iS_1$. Thus necessarily $R_1 = R$ and $S_1 = S$. Hence the <u>Hermitian</u> <u>decomposition</u> $T = R + iS$ of T is uniquely determined. It is an interesting fact (noted in Halmos [9], p. 42) that an operator T is normal if and only if $RS = SR$, where $T = R + iS$ is the Hermitian decomposition of T.

We conclude this chapter with a theorem on the invertibility of operators of the form $I + T^*T$, which will be required for a later result.

<u>Theorem 4</u>. Let T be an operator on H. Then the operators T^*T and TT^* are Hermitian and $I + T^*T$ and $I + TT^*$ are both invertible operators.

The above theorem is due to J. von Neumann [21]. A proof of a more general theorem is given by Yosida, (Yosida [30], p. 200).

CHAPTER III

BACKGROUND IN THEORY OF BANACH ALGEBRAS

1. Complex Commutative Algebras

Definition 1. A linear space A over the complex field C is a complex commutative algebra if there is a multiplication defined which makes correspond to each pair $x,y \in A$ and element $xy \in A$, called the vector product of x and y, which satisfies the following conditions for all $x,y,z \in A$ and all $\lambda \in C$:

$$\lambda(xy) = (\lambda x)y = x(\lambda y) \tag{1}$$

$$(x + y)z = xz + yz \tag{2}$$

$$xy = yx \ . \tag{3}$$

The term commutative algebra will be used to refer to a complex commutative algebra. A commutative algebra A is said to have the unit (or identity) element e, if A contains an element e such that $xe = x$ for every $x \in A$. The terminology is justified, since if $xe = x$ and $xe' = x$ for all $x \in A$, then $e = ee' = e'e = e'$, and consequently there can be at most one unit element. Every commutative algebra to be used in the following discussion will be a commutative algebra with the unit element e. If A is such an algebra, and $x \in A$, x is called invertible if there exists an element $x^* \in A$ such that $xx^* = e$. If $y^* \in A$ is such $xy^* = e$, then

$$x^* = x^*e = x^*(xy^*) = (x^*x)y^* = ey^* = y^* .$$

Hence, if x is invertible, there is precisely one element in A to be denoted x^{-1}, such that $xx^{-1} = e$. The element $x^{-1} \in A$ is called the <u>inverse of x</u>. An element $x \in A$ which is not invertible is said to be <u>singular</u>.

<u>Definition 2</u>. Let $x \in A$, where A is a commutative algebra with unit e. The set of all complex numbers λ such that $x - \lambda e$ is not invertible is called the <u>spectrum</u> of x, and is denoted by $\sigma(x)$. The complement of $\sigma(x)$ in C is called the <u>resolvent set</u> of x, denoted by $r(x)$. (Note that this terminology corresponds to that used in Chapter II, Section 2.)

<u>Definition 3</u>. Let A be a commutative algebra with unit. A subset $J \subset A$ is called an <u>ideal</u> of A if

$$J \text{ is a linear subspace of A, and} \tag{4}$$

$$x \in J \text{ and } y \in A \text{ implies } xy \in J . \tag{5}$$

It is clear that A necessarily has the two ideals $\{0\}$ and A itself. If J is an ideal of A, and $J \neq A$, then we say that J is a <u>proper ideal</u> of A. It should be noted here that a proper ideal J can contain no invertible element x, for if so $e = x^{-1}x \in J$, and it would follow that $J = A$. A <u>maximal ideal</u> J of A is a proper ideal of A which is not properly contained in any proper ideal of A.

Theorem 1. Let A be a commutative algebra with unit e. Every proper ideal of A is contained in a maximal ideal of A.

Proof. Suppose S is the set of all proper ideals of A which contain the proper ideal J. Since $J \in S$, S is nonempty. Let S be partially ordered by set inclusion. Let $\{J_t : t \in T\}$ be a chain in S, that is, a totally ordered subset of S, and define $K = \cup_{t \in T} J_t$. We show that K is an ideal of A, and that K is a proper ideal. Suppose $x,y \in K$, $z \in A$, and $\alpha,\beta \in C$. There exist indices s and t in T such that $x \in J_s$, $y \in J_t$. Since $\{J_t : t \in T\}$ is a chain in S, either $J_s \subset J_t$ or $J_t \subset J_s$. Suppose $J_s \subset J_t$. Then both x and y belong to the ideal J_t, and thus $\alpha x + \beta y \in J_t \subset K$ and $xz \in J_t \subset K$. A similar argument shows that $\alpha x + \beta y \in K$ and $xz \in K$ if $J_t \subset J_s$. Hence K is an ideal of A. Since each J_t is proper, no J_t contains the identity e of A, and thus $e \notin K$. Hence K is a proper ideal of A, and $K \supset J$ since each $J_t \supset J$. Consequently K is an upper bound in S for the chain. Zorn's lemma asserts that S has a maximal element M. Such an element $M \in S$ is a maximal ideal of A containing J.

Corollary 1. If A is as in Theorem 1, and $x \in A$ is singular (non-invertible), there exists a maximal ideal of A containing the element x.

Proof. Let $J_x = \{xy : y \in A\}$. It is easily verified that J_x is an ideal of A containing x. If $e \in J_x$, there would exist an element $y \in A$ such that $xy = e$. But this would contradict the hypothesis that x is singular. Hence $e \notin J_x$, and thus J_x is a proper ideal of A. By Theorem III.1.1, there is a maximal ideal of A containing J_x and thus containing the singular element x.

Note. It follows from Corollary 1 and an earlier remark that an element $x \in A$ is invertible if and only if x is contained in no maximal ideal of A.

Definition 4. Let X and X' be commutative algebras. A mapping $\Lambda : X \to X'$ is said to be a homomorphism of X into X' if for all $x,y \in X$ and all $\alpha,\beta \in C$

$$\Lambda(\alpha x + \beta y) = \alpha\Lambda(x) + \beta\Lambda(x) \tag{6}$$

$$\Lambda(xy) = \Lambda(x)\Lambda(y) \quad . \tag{7}$$

An isomorphism is a one-to-one homomorphism. The inverse image of the zero element of X' is called the kernel of the homomorphism Λ.

Let J be a proper ideal of A, where A is a commutative algebra with unit. The relation $\sim$ will be defined as follows. For $x,y \in A$, $x \sim y$ (modulo J) if and only if $x - y \in J$. This relation is easily shown to be an equivalence relation in A, and A is thereby partitioned into disjoint equivalence classes. The set $\{y : x-y \in J\} \subset A$ is called the residue class (modulo J) containing x, and will be denoted by $\varphi(x)$. Note that $\varphi(x)$ is precisely the coset of J defined by

$$x + J = \{x + z : z \in J\} \quad . \tag{8}$$

For all $x,y \in A$, and all scalars λ, define $\lambda\varphi(x)$, $\varphi(x) + \varphi(y)$ and $\varphi(x)\varphi(y)$ by

$$\lambda\varphi(x) = \varphi(\lambda x) \quad , \tag{9}$$

$$\varphi(x) + \varphi(y) = \varphi(x + y), \tag{10}$$

$$\varphi(x)\varphi(y) = \varphi(xy) \quad . \tag{11}$$

It is easily verified that $\lambda\varphi(x)$, $\varphi(x) + \varphi(y)$ and $\varphi(x)\varphi(y)$ are residue classes (modulo J), and that the operations above are well-defined. Consider, for example, the proposed definition of the product $\varphi(x)\varphi(y)$ of the coset $\varphi(x)$ and $\varphi(y)$. Suppose $x_1 \in \varphi(x)$, $y_1 \in \varphi(y)$. Then $x_1 - x \in J$ and $y_1 - y \in J$. Since J is an ideal, and

$$x_1 y_1 - xy = x_1(y_1 - y) + (x_1 - x)y \quad ,$$

it follows that $x_1 y_1 - xy \in J$. Thus $\varphi(x_1 y_1) = \varphi(xy)$. Consequently, $\varphi(xy)$ is determined solely by the <u>cosets</u> $\varphi(x)$ and $\varphi(y)$, and the definition $\varphi(x)\varphi(y) = \varphi(xy)$ is unambiguous. The set of all residue classes (modulo J), denoted by A/J, is called the <u>residue class algebra</u> (or <u>quotient space</u>) of A (modulo J). This terminology is appropriate since the set A/J becomes a commutative algebra with unit, when endowed with the operations defined by (9), (10) and (11), and the unit element $e + J = \varphi(e)$.

The mapping $\varphi : A \to A/J$ which assigns the coset $\varphi(x)$ to each

$x \in A$ is clearly a homomorphism by definition of the operations in A/J. This mapping φ is known as the natural homomorphism of A onto A/J. The image $\varphi(e)$ of the unit $e \in A$ is the unit element of A/J, and $\varphi(0)$ is the zero element of A/J.

Any homomorphism $\Lambda : X \to X'$, where X and X' are commutative algebras, can be described in terms of ideals in the domain space, as indicated in the following theorem.

Theorem 2. Under a homomorphism $\Lambda : X \to X'$ the kernal K of Λ is an ideal of X. The image $\Lambda(X)$ of X is itself isomorphic to the residue class algebra X/K. Conversely, every ideal J of X is the kernal of the natural homomorphism of X into X/J.

Refer, for example, to Naimark [19], p. 166 for a proof of this theorem.

Definition 5. A commutative algebra A with unit is called simple if it contains no proper ideals different from the trivial ideal $\{0\}$.

Maximal ideals play an important role here.

Theorem 3. The residue class algebra A/J is simple if and only if J is a maximal ideal.

A proof is given in Naimark [19], p. 167.

Definition 6. Let M denote the set of all maximal ideals of A. The radical R of A is the ideal $\cap_{J \in M} J$. A is said to be semi-simple if $R = \{0\}$.

Note that every simple algebra is also semi-simple. The algebra of complex numbers is a simple algebra in the sense of Definition III.1.5.

A condition under which the residue class algebra A/J is also a field is given in the following theorem.

Theorem 4. Let J be a maximal ideal of a commutative algebra A with unit e. Then the commutative algebra A/J is a field.

Proof. It is only necessary to show that each nonzero element of A/J has a multiplicative inverse. This will imply that the nonzero elements of A/J form a group under multiplication. Suppose $x \in A$, but $x \notin J$. Then $\varphi(x) \neq \varphi(0)$, and $\varphi(x)$ is a nonzero element of A/J. (Note that such an element x exists since J is a proper ideal of A.) Define

$$\check{J} = \{ax + b : a \in A,\ b \in J\} \ .$$

It is easily verified that $\check{J}$ is an ideal of A containing J, and $x \in \check{J}$ since $x = ex + 0$. Thus $\check{J}$ contains J properly. Since J is a maximal ideal, it follows that $\check{J} = A$. Since $\check{J} = A$, there exists elements $a \in A$, $b \in J$ such that $ax + b = e$. It follows that $ax = e + (-b) \in \varphi(e)$. Thus $\varphi(ax) = \varphi(e)$, and $\varphi(a)\varphi(x) = \varphi(e)$. Hence $\varphi(a)$ is the multiplicative inverse of $\varphi(x)$. This completes the proof.

2. Abstract Analytic Functions

In this section we indicate how the theory of analytic functions can be generalized to functions of a complex variable with values in a Banach space. The intent here is merely to indicate, without detailed arguments, how the generalization is made. Specific references are supplied.

Definition 1. A linear space X over the complex field C is a normed linear space if there is a mapping which makes correspond to each

$x \in X$ a real number $\|x\|$, called the <u>norm</u> of x, which satisfies the following conditions for all $x,y \in X$ and every $\lambda \in C$:

$$\|x\| \geq 0 \text{ and } \|x\| = 0 \text{ if and only if } x = 0 \tag{1}$$

$$\|\lambda x\| = |\lambda| \cdot \|x\| \tag{2}$$

$$\|x + y\| \leq \|x\| + \|y\| \quad . \tag{3}$$

A metric topology in X is obtained by defining the metric d by $d(x,y) = \|x - y\|$ for every $x,y \in X$. The open sets of this topology are specified in precisely the same manner as in the metric topology in a pre-Hilbert space discussed in Section 1 of Chapter II. A normed linear space which is <u>complete</u> as a metric space (with the metric defined above) is called a <u>Banach space</u>.

<u>Definition 2</u>. Let $\{x_n\}$ be a sequence in a normed linear space X. If there exists a point $x \in X$ such that

$$\lim_{n \to \infty} \|x_n - x\| = 0,$$

then x will be called the <u>strong limit</u> of the sequence $\{x_n\}$, and we write

$$x = s - \lim_{n \to \infty} x_n \quad .$$

Definition 3. Let $x(\cdot)$ be a function defined on a subset D of the complex plane C with values in a normed linear space X. Let λ_o be a limit point of D, and suppose that there exists a point $y \in X$ such that the following condition is satisfied: For every open set $S \subset X$ containing y, there exists an open set $T \subset C$ containing λ_o, such that

$$\lambda \in (T - \{\lambda_o\}) \cap D \text{ implies } x(\lambda) \in S \quad .$$

Then we will say that y is the strong limit of $x(\cdot)$ as λ approaches λ_o through S, and we shall write

$$y = s - \lim_{\lambda \to \lambda_o} x(\lambda) \quad .$$

If $x(\cdot)$ is defined at λ_o, and

$$s - \lim_{\lambda \to \lambda_o} x(\lambda) = x(\lambda_o) \quad ,$$

then $x(\lambda)$ is said to be strongly continuous at λ_o. If $x(\cdot)$ is strongly continuous at every point of S, then $x(\cdot)$ is said to be strongly continuous on S.

Definition 4. Let D be a region of the complex plane, and X a normed linear space. We will say that $x(\cdot):D \to X$ is strongly differentiable at the point $\lambda_o \in D$ if the strong limit as λ approaches λ_o of the quotient

$$\frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o}$$

exists. In this case we shall write

$$x'(\lambda_o) = s - \lim_{\lambda \to \lambda_o} \left\{ \frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o} \right\} \tag{4}$$

and call $x'(\lambda_o)$ the <u>derivative</u> of $x(\cdot)$ at λ_o. If there exists a neighborhood $N(\lambda_o;\varepsilon) = \{\lambda : \lambda \in C, |\lambda - \lambda_o| < \varepsilon\}$ of $\lambda_o \in D$ such that $x(\cdot)$ is strongly differentiable at every point $\lambda \in N(\lambda_o;\varepsilon)$, then $x(\cdot)$ is said to be <u>strongly holomorphic</u> (or analytic) <u>at λ_o</u>. The function $x(\cdot)$ is said to be <u>strongly holomorphic in D</u> if $x(\cdot)$ is strongly holomorphic at every point $\lambda \in D$.

<u>Note</u>. The usual notation will be adopted for "n^{th} derivatives" of a function $x(\cdot): C \to X$. Thus, assuming that the derivatives $x'(\lambda)$, $(x')'(\lambda)$ and so on, exist for $\lambda \in D$, we write $x^{(1)}(\lambda) = x'(\lambda)$, $x^{(2)}(\lambda) = (x')'(\lambda)$, ... We define, for convenience, $x^{(0)}(\lambda) = x(\lambda)$. For any positive integer n, $x^{(n)}(\lambda)$ will be called the <u>n^{th} derivative of $x(\cdot)$ at λ</u>.

In the following we denote by X^* the space of all bounded linear functionals on the normed linear space X, so that X^* is the <u>dual space</u> of X.

<u>Lemma 1</u>. Let D be a region in C, and let X be a normed linear space. Suppose $x(\cdot): D \to X$ is strongly holomorphic in D, and let $f \in X^*$. Then the composite function $f \cdot x$, defined by $(f \cdot x)(\lambda) = f[x(\lambda)]$ for every $\lambda \in D$, is a complex-valued holomorphic function on D.

Proof. Let $\lambda_o \in D$. Since $x(\cdot)$ is strongly holomorphic in D,

$$x'(\lambda_o) = s - \lim_{\lambda \to \lambda_o} \left\{ \frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o} \right\}$$

exists. By linearity of f it follows that, for every $\lambda \in D$,

$$\frac{f[x(\lambda)] - f[x(\lambda_o)]}{\lambda - \lambda_o} = f\left[\frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o} \right] .$$

The bounded linear functional f is necessarily a continuous function. Hence

$$(f \cdot x)'(\lambda_o) = \lim_{\lambda \to \lambda_o} \left\{ \frac{f[x(\lambda)] - f[x(\lambda_o)]}{\lambda - \lambda_o} \right\} = f\left\{ s - \lim_{\lambda \to \lambda_o} \left[\frac{x(\lambda) - x(\lambda_o)}{\lambda - \lambda_o} \right] \right\}$$

$$= f[x'(\lambda_o)] \quad . \tag{5}$$

Since $\lambda_o \in D$ was arbitrarily chosen, (5) implies that $f \cdot x$ is holomorphic on D.

Lemma 2. Let X be a normed linear space, and let $x \in X$ be fixed. If $f(x) = 0$ for every $f \in X^*$, then $x = 0$.

The above lemma is an easy consequence of the Hahn-Banach theorem. For a proof of this result, reference may be made to Bachman and Narici [1], p. 199.

Theorem 1. (An Abstract Liouville Theorem). Let X be a normed linear space, and suppose $x(\cdot): C \to X$ is strongly holomorphic throughout C. If there exists a real number $M > 0$ such that

$$\|x(\lambda)\| < M \text{ for all } \lambda \in C,$$

then $x(\cdot)$ must be constant on C.

<u>Proof</u>. Let $f \in X^*$ be any bounded linear functional on X. Then $f \cdot x$ is holomorphic on C by Lemma 1. Moreover, we have

$$\|f[x(\lambda)]\| \leq \|f\| \cdot \|x(\lambda)\| \leq \|f\| \cdot M$$

for every $\lambda \in C$. Therefore $f \cdot x$ is a bounded function which is holomorphic throughout C. By the ordinary Liouville Theorem (see Markushevich [16], p. 364), the complex-valued function $f \cdot x$ must be constant on C. Let μ and ξ be any two complex numbers. By the above remarks it follows that

$$f[x(\mu)] = f[x(\xi)] .$$

By linearity of $f \in X^*$, this implies

$$f[x(\mu) - x(\xi)] = 0.$$

But f was any bounded linear functional on X, hence

$$x(\mu) - x(\xi) = 0 , \qquad (6)$$

by Lemma 2. The relation (6) asserts that $x(\cdot)$ is constant on C.

Note. The above proof follows the general lines of that in Bachman and Narici [1].

Definition 5. By a contour Γ connecting the points γ_1 and γ_2 in C, we will mean a rectifiable Jordan arc in C from γ_1 to γ_2. If Γ is a closed Jordan curve ($\gamma_1 = \gamma_2$), then Γ is called a closed contour in C.

Note. Every contour Γ has a parametrization $\gamma : [t_1,t_2] \to \Gamma$, where $\gamma(t) = \alpha(t) + i\beta(t)$ for $t_1 \leq t \leq t_2$ and $\gamma_1 = \gamma(t_1)$, $\gamma_2 = \gamma(t_2)$. The mapping γ is continuous and one-to-one on $[t_1,t_2)$. If Γ is a closed contour, then $\gamma(t_1) = \gamma(t_2)$. Γ is to be thought of as an oriented curve in the following sense: if $\lambda', \lambda'' \in \Gamma$ and $\lambda' = \gamma(t') \neq \gamma(t'') = \lambda''$, then λ' is said to precede λ'' if $t' < t''$. A point $\lambda_o \in \Gamma$ is said to be between the points $\lambda', \lambda'' \in \Gamma$ if $\lambda_o = \gamma(t_o)$, $\lambda' = \gamma(t')$ and $\lambda'' = \gamma(t'')$ implies $t' < t_o < t''$.

According to the Jordan curve theorem (see Markushevich [16], p. 70), the complement Γ^c of any closed Jordan curve Γ in C has exactly two components with Γ as their common boundary. The bounded component of Γ^c will be called as usual the interior of Γ, denoted $I(\Gamma)$, and the unbounded component of Γ^c will be called the exterior of Γ, denoted $E(\Gamma)$.

Definition 6. Let Γ be a contour connecting the points γ_1 and γ_2 in C. Let X be a normed linear space, and $x(\cdot)$: C $\to$ X. Let Π denote a partition $\gamma_1 = \lambda_o, \lambda_1, \ldots, \lambda_n = \gamma_2$ of Γ (where $\lambda_i \neq \lambda_j$ if $i \neq j$), and denote by Ω a set of points $\xi_1, \xi_2, \ldots, \xi_n \in \Gamma$ which are between the points of Π. For the partition Π, let

$$\|\Pi\| = \max\{|\lambda_i - \lambda_{i-1}| : i = 1, 2, \ldots, n\} ,$$

and denote by $S(\Pi, \Omega)$, the sum

$$S(\Pi,\Omega) = \sum_{i=1}^{n} x(\xi_i)\cdot(\lambda_i - \lambda_{i-1}) \in X. \quad (7)$$

We will say that <u>the strong limit</u>

$$\int_\Gamma x(\lambda)d\lambda = \underset{\substack{\|\Pi\| \to 0 \\ n \to \infty}}{s - \lim} S(\Pi,\Omega) \quad (8)$$

<u>exists</u> if the following condition is satisfied: for every $\epsilon > 0$ there exists a number $\delta > 0$ such that for every partition Π of Γ with $\|\Pi\| < \delta$ and for every choice of Ω, we have

$$\|S(\Pi,\Omega) - \int_\Gamma x(\lambda)d\lambda\| < \epsilon . \quad (9)$$

If the above condition holds (that is, if the strong limit (8) exists), then we say that $x(\cdot)$ is <u>integrable along</u> Γ and the strong limit $\int_\Gamma x(\lambda)d\lambda$ is called the <u>contour integral</u> of $x(\cdot)$ along Γ.

<u>Theorem 2</u>. Let Γ be a contour in C. Let X be a Banach space, and suppose that $x(\cdot):C \to X$ is strongly continuous at every point of Γ. Then $x(\cdot)$ is integrable along Γ.

The above theorem can be proved as a consequence of the completeness of X, and the uniform continuity of $x(\cdot)$ on the compact set Γ in C. A proof of this existence theorem is given in a more general setting in Hille and Phillips [11] p. 63.

Theorem 3. (An Abstract Cauchy Integral Theorem). Let X be a Banach space. Suppose that $x(\cdot)C \to X$ is strongly holomorphic on a region D bounded by the contour Γ and that $x(\cdot)$ is strongly continuous on Γ. Then

$$\int_\Gamma x(\lambda)d\lambda = 0 .$$

Proof. Let $f \in X^*$, and let $y = \int_\Gamma x(\lambda)d\lambda \in X$. Since f is a bounded linear functional on X, we have

$$f(y) = f[\int_\Gamma x(\lambda)d\lambda] = \int_\Gamma f[x(\lambda)]d\lambda . \tag{10}$$

Relation (10) is a consequence of the linearity and continuity of f on X. By Lemma III.2.1, $f \cdot x$ is a complex-valued holomorphic function on D. Also, $f \cdot x$ is easily seen to be continuous on the contour Γ which bounds D. By the ordinary Cauchy integral theorem (see Markushevich [16], p. 258), we have

$$\int_\Gamma f[x(\lambda)]d\lambda = 0 = f(y) .$$

Since $f \in X^*$ was arbitrary, it follows by Lemma III.2.2 that $y = 0$.

Note. The above proof is similar to that in Bachman and Narici [1], p. 310. The other well-known results of ordinary analytic function theory carry over to the theory of holomorphic functions with values in a Banach space. We can use Lemmas III.2.1 and III.2.2 in the manner typified in the above proof to obtain the more general or "abstract" versions of the Cauchy integral formula and the Taylor series expansion for holomorphic functions. For the proofs of these theorems for ordinary analytic functions, reference may be made to Markushevich [16]. We omit the proofs of the following theorems.

Theorem 4. (An Abstract Cauchy Integral Formula). Let X be a Banach space, and suppose $x(\cdot):C \to X$ is strongly holomorphic in a region D of C which contains a closed contour Γ and the interior of Γ, $I(\Gamma)$. Then

$$\frac{1}{2\pi i}\int_{\Gamma} \frac{x(\xi)}{\xi - \lambda}\, d\xi = x(\lambda) \text{ if } \lambda \in I(\Gamma) \quad ,$$

and

$$\frac{1}{2\pi i}\int_{\Gamma} \frac{x(\xi)}{\xi - \lambda}\, d\xi = 0 \text{ if } \lambda \in E(\Gamma) \quad .$$

Note. Under the hypotheses of Theorem 4, we can also prove the formula

$$x^{(n)}(\lambda) = \frac{n!}{2\pi i}\int_{\Gamma} \frac{x(\xi)}{(\xi - \lambda)^{n+1}}\, d\xi \tag{11}$$

for the n^{th} derivative $x^{(n)}$ of $x(\cdot)$ at $\lambda \in I(\Gamma)$.

Theorem 5. (An Abstract Taylor Series Expansion). Let X be a Banach space, and suppose $x(\cdot):C \to X$ is strongly holomorphic in a region D of C. Let λ_o be an arbitrary point of D. Then $x(\cdot)$ is represented by the power series

$$x(\lambda) = \sum_{n=0}^{\infty} \frac{x^{(n)}(\lambda_o)}{n!} (\lambda - \lambda_o)^n \quad , \tag{12}$$

which converges on the disk $|\lambda - \lambda_o| < \delta$, where $\delta = \inf\{|\lambda_o - \mu| : \mu \in \partial D\}$ is the distance between λ_o and ∂D, the boundary of D. The n^{th} derivative $x^{(n)}(\lambda_o)$ is given by

$$x^{(n)}(\lambda_o) = \frac{n!}{2\pi i} \int_\Gamma \frac{x(\lambda)}{(\lambda - \lambda_o)^{n+1}} d\lambda \quad ,$$

for n = 1, 2, 3, ...

Note. If $x(\cdot)$ is strongly holomorphic on the entire complex plane C, then $\delta = \infty$, and the power series (12) converges for all $\lambda \in C$.

3. Banach Algebras

Definition 1. A complex commutative algebra A with unit e is a commutative normed algebra if A is a normed linear space by a norm $\|\cdot\|$ which also satisfies the following conditions:

$$\|xy\| \leq \|x\| \cdot \|y\| \quad \text{for all } x,y \in A \tag{1}$$

$$\|e\| = 1 \quad . \tag{2}$$

Only commutative normed algebras A will be considered in what follows.

A commutative normed algebra which is a Banach space is called a commutative Banach algebra. The complex field itself is a commutative

Banach algebra with the norm $\|\lambda\| = |\lambda|$, for every complex number λ.

Theorem 1. Every maximal ideal J of a commutative Banach algebra A (with unit e) is a closed set in the normed linear space A.

Proof. Suppose $x, y \in \bar{J}$ (the closure of J in A), and $\alpha, \beta \in C$. Let $\{x_n\}$, $\{y_n\}$ be sequences of elements of A such that $x_n \to x$ and $y_n \to y$. Then $\alpha x_n + \beta y_n \to \alpha x + \beta y$. Since $\alpha x_n + \beta y_n \in J$ for every n, $\alpha x + \beta y \in \bar{J}$. If $z \in A$, then $zx_n \in J$ for every n. Since $zx_n \to zx$, it follows that $zx \in \bar{J}$. Hence $\bar{J}$ is an ideal of A.

Since J is a proper ideal (as is every maximal ideal, by definition), no invertible element of A belongs to J. Thus all elements of J are singular elements of A. It is established in lemmas following this argument that the set G of all singular elements of A is a closed set in A. Assuming this result, it follows from $J \subset G$ that $\bar{J} \subset \bar{G} \subset G$. Now G is a proper subset of A (since $e \notin G$). Thus $\bar{J}$ is properly contained in A, and hence is a proper ideal of A. Since J is a maximal ideal of A, and $J \subset \bar{J}$, then necessarily $J = \bar{J}$. Consequently J is closed in A.

Lemma 1. If A is a commutative Banach algebra with unit e, $x \in A$, and $\|x\| < 1$, then $e + x$ is invertible.

Proof. By a property of the norm in A, $\|x^n\| \leq \|x\|^n$ for n = 1, 2, 3, ... Let

$$y_n = e - x + x^2 - \ldots + (-1)^n x^n ,$$

and note that each $y_n \in A$. If $k \geq 1$,

$$\|y_{n+k} - y_n\| \le \sum_{j=n+1}^{n+k} \|x\|^j \le \frac{\|x\|^{n+1}}{1 - \|x\|} .$$

This estimate shows that $\{y_n\}$ is a Cauchy sequence in A. Since A is complete as a normed linear space, there exists a $y \in A$ such that $y_n \to y$ (in the norm of A). The identity

$$(e + x)y_n = e + (-1)^{n+1}x^{n+1}$$

yields, upon taking limits as $n \to \infty$,

$$(e + x)y = e .$$

Thus e + x is invertible, with

$$(e + x)^{-1} = y = e + \sum_{n=1}^{\infty} (-1)^n x^n . \qquad (3)$$

The series (3) is often called the <u>Neumann series</u> for $(e + x)^{-1}$.

<u>Corollary 1</u>. Under the hypotheses of Lemma 1,

$$\|(e + x)^{-1} - e\| \le \frac{\|x\|}{1 - \|x\|} .$$

<u>Lemma 2</u>. If A is a commutative Banach algebra with identity e, if x is an invertible element of A, and $h \in A$ is such that $\|h\| < 1/\|x^{-1}\|$, then the element x + h is invertible. The set G of invertible elements of A is thus open in A, and the complementary set G^c, the set of

singular elements of A, is closed in A.

<u>Proof</u>. Note that $x + h = x(e + x^{-1}h)$. By a property of the norm,

$$\|x^{-1}h\| \leq \|x^{-1}\| \cdot \|h\| < 1$$

since $\|h\| < 1/\|x^{-1}\|$. By Lemma 1, $e + x^{-1}h$ is invertible. Hence

$$(x + h)\ (e + x^{-1}h)^{-1}x^{-1} = e.$$

Hence $x + h$ is invertible, with

$$(x + h)^{-1} = (e + x^{-1}h)^{-1}x^{-1} .$$

It has been proved that if x is invertible, every element of A in some open ball about x is invertible. Thus the set of invertible elements of A is open in A.

<u>Corollary 2</u>. Under the hypotheses of Lemma 2, the mapping $x \to x^{-1}$ defined on the set G of invertible elements is continuous.

<u>Proof</u>. Note that

$$(x + h)^{-1} - x^{-1} = [(e + x^{-1}h)^{-1} - e]x^{-1} .$$

By the Corollary of Lemma 1, it follows that

$$\|(x + h)^{-1} - x^{-1}\| \leq \|(e + x^{-1}h)^{-1} - e\| \cdot \|x^{-1}\|$$

$$\leq \frac{\|x^{-1}h\|}{1 - \|x^{-1}h\|} \|x^{-1}\| \leq \frac{\|x^{-1}\|^2 \cdot \|h\|}{1 - \|x^{-1}\| \cdot \|h\|} \quad .$$

This inequality implies the assertion of the corollary.

Note. The proofs of Theorem III.3.1 and the lemmas are modified versions of similar proofs in Rudin [25], pp. 352-353 and p. 357.

Definition 2. Let $x \in A$, where A is a commutative Banach algebra with unit e. The spectral radius of x, denoted $\rho(x)$, is the radius of the smallest closed disk in C with center at the origin which contains the spectrum $\sigma(x)$; in symbols,

$$\rho(x) \quad = \quad \sup\{|\lambda| \;:\; \lambda \in \sigma(x)\} \quad . \tag{4}$$

It can be shown that for $x \in A$, the spectrum of x is a compact set in the complex plane (see Rickart [22], p. 30). In view of this result the complement of $\sigma(x)$ in C, the resolvent set $r(x)$, is an open unbounded subset of C.

Definition 3. The resolvent function of an element $x \in A$ is the function $R_x : C \to A$ defined by

$$R_x(\lambda) \quad = \quad (x - \lambda e)^{-1} \text{ for every } \lambda \in r(x) \quad . \tag{5}$$

Remark 1. The resolvent function R_x of $x \in A$ satisfies the relation

$$R_x(\xi) - R_x(\eta) \quad = \quad (\xi - \eta)\, R_x(\xi)\, R_x(\eta) \quad . \tag{6}$$

To deduce this formula, note that

$$(x - \eta e)(x - \xi e)^{-1} = (x - \xi e + \xi e - \eta e)(x - \xi e)^{-1}$$

$$= e + (\xi - \eta)(x - \xi e)^{-1} .$$

Upon multiplication by $(x - \eta e)^{-1}$, we obtain

$$(x - \xi e)^{-1} = (x - \eta e)^{-1} + (\xi - \eta)(x - \xi e)^{-1}(x - \eta e)^{-1} ,$$

which is the formula (6). Fix $\xi \in r(x)$. Then for $\eta \in r(x)$, $\eta \neq \xi$, it follows from (6) that

$$\frac{R_x(\eta) - R_x(\xi)}{\eta - \xi} = R_x(\xi) \cdot R_x(\eta) = (x - \xi e)^{-1}(x - \eta e)^{-1} .$$

It has been proved (Corollary to Lemma III.3.2) that the mapping $x \to x^{-1}$ is continuous on the set G of invertible elements of A. Hence the strong limit

$$s - \lim_{\eta \to \xi} \frac{R_x(\eta) - R_x(\xi)}{\eta - \xi} = (x - \xi e)^{-1}(x - \xi e)^{-1} = (R_x(\xi))^2 \in A .$$

This calculation shows that R_x is a strongly holomorphic function on the set r(x).

<u>Theorem 2</u>. Let A be a commutative Banach algebra with unit e. For every element $x \in A$, $\sigma(x)$ is nonempty.

Proof. Suppose, on the contrary, that for some $x \in A$, $\sigma(x) = \phi$. We will show that the above assumption gives a contradiction. It has been shown (in Remark 1) that the resolvent function R_x of x is strongly holomorphic on the resolvent set $r(x) = \sigma(x)^c$. Since $\sigma(x)$ is empty, R_x is strongly holomorphic everywhere in C. We also have

$$\|(e - \tfrac{x}{\lambda})-e\| = \|\tfrac{x}{\lambda}\| = \frac{1}{|\lambda|}\|x\| \to 0 \text{ as } |\lambda| \to \infty ,$$

whereby

$$(e - \tfrac{x}{\lambda}) \to e \text{ in norm as } |\lambda| \to \infty .$$

Since the mapping $x \to x^{-1}$ is a continuous function on the set G of invertible elements of A (Corollary to Lemma III.3.2), it follows that

$$(e - \tfrac{x}{\lambda})^{-1} \to e^{-1} = e \text{ in norm as } |x| \to \infty . \tag{7}$$

By (7) we obtain

$$\|(x - \lambda e)^{-1}\| = |\lambda^{-1}|\cdot\|(e - \tfrac{x}{\lambda})^{-1}\| \to \frac{\|e\|}{|\lambda|} \to 0 \tag{8}$$

as $|\lambda| \to \infty$. Hence for any given $\varepsilon_1 > 0$, there exists a positive real δ_1 such that $\|(x - \lambda e)^{-1}\| < \varepsilon_1$ for $|\lambda| > \delta_1$. Let

$$\varepsilon_2 = \max_{|\lambda|\le\delta_1} \|(x - \lambda e)^{-1}\| ,$$

and let $M = \max \{\epsilon_1, \epsilon_2\}$. Then

$$\|R_x(\lambda)\| = \|(x - \lambda e)^{-1}\| \leq M \text{ for every } \lambda \in C.$$

By Theorem 1 of Section 2, R_x must be constant on C, since R_x is strongly holomorphic everywhere in C and is bounded on C. Thus

$$R_x(\lambda) = x_o \in A \text{ for every } \lambda \in C .$$

But (8) implies that $x_o = 0$, which contradicts the fact that

$$(x - \lambda e)(x - \lambda e)^{-1} = e.$$

Therefore, we must have $\sigma(x) \neq \emptyset$ for every $x \in A$. This completes the proof..

<u>Definition 4</u>. A mapping $\Lambda: X \to X'$ of a normed space X into a normed space X' is called an <u>isometry</u> if Λ preserves norms; that is, $\|\Lambda(x)\| = \|x\|$ for every $x \in X$. If the spaces X and X' are commutative normed algebras and Λ is an isomorphism as well as an isometry, then we say that Λ is an <u>isometric isomorphism</u>.

The following theorem plays a very important role in the theory of Banach algebras.

<u>Theorem 3</u> (<u>Gelfand-Mazur</u>). If A is a commutative Banach algebra with unit e in which every non-zero element is invertible, then A is isometrically isomorphic to the field C of complex numbers. This isomorphism occurs in the following manner. Every element $x \in A$ is of

the form $x = \lambda e$, where $\lambda \in C$.

Proof. Let $x \in A$. By Theorem III.3.2, $\sigma(x) \neq \emptyset$. Hence there is at least one complex number $\lambda \in \sigma(x)$, and $x - \lambda e$ is not invertible. Since, by hypothesis, every non-zero element of A is invertible, we must have

$$x - \lambda e = 0 .$$

Therefore $x = \lambda e$. Denote the complex number λ, which depends on $x \in A$, by $\lambda(x)$. Then the mapping $x \to \lambda(x)$ is easily seen to be an isomorphism from A onto C. For example, if $x,y \in A$, then $x + y = \lambda(x)e + \lambda(y)e = [\lambda(x) + \lambda(y)]e$ and $x + y = \lambda(x + y)e$. Thus, $\lambda(x + y) = \lambda(x) + \lambda(y)$. Similarly, the manner in which $x \to \lambda(x)$ is defined also gives

$$\lambda(\alpha x) = \alpha\lambda(x) \quad \text{and} \quad \lambda(xy) = \lambda(x)\lambda(y) \quad ,$$

where $x,y \in A$ and $\alpha \in C$. Clearly $x \to \lambda(x)$ is one-to-one since $\lambda(x) = \lambda(y)$ gives

$$x = \lambda(x)e = \lambda(y)e = y \quad .$$

Therefore $x \to \lambda(x)$ is an isomorphism. For every $x \in A$,

$$|\lambda(x)| = |\lambda(x)| \cdot \|e\| = \|\lambda(x)e\| = \|x\| \quad .$$

Hence $x \to \lambda(x)$ is also an isometry.

Note. The proof that for every $x \in A$, $x = \lambda e$ for some $\lambda \in C$, is similar to that given in Gelfand, Raikov and Shilov [8], p. 31.

Definition 5. Let J be a closed ideal of A. For every element $\varphi(x) \in A/J$ define the quotient norm of $\varphi(x)$ by

$$\|\varphi(x)\| = \inf\{\|y\| : y \in \varphi(x)\} \quad . \tag{9}$$

Note that the quotient norm of $\varphi(x) \in A/J$ is also given by

$$\|\varphi(x)\| = \inf\{\|x + y\| : y \in J\} \tag{10}$$

(see Rudin [25], p. 358), called the distance from x to J.

Theorem 4. Let J be a closed ideal of A. Then:

A/J is a normed algebra by the quotient norm. (11)

If A is a commutative Banach algebra and J is a proper closed ideal, then A/J is a commutative Banach algebra. (12)

For a proof, refer to Rudin [25], p. 359.

Corollary 3. If J is a maximal ideal of the commutative Banach algebra A with unit, then A/J is a commutative Banach algebra with unit in which every non-zero element has an inverse.

Proof. The corollary follows directly from Theorems III.1.4, III.3.4, and the fact that every maximal ideal in the Banach algebra A is a closed set in A, by Theorem III.3.1.

It is now possible to characterize the elements of the residue class algebra A/J, where J is a maximal ideal of the commutative Banach algebra A with unit, by applying the Gelfand-Mazur Theorem.

Theorem 5. Let A be a commutative Banach algebra with unit, and let J be a maximal ideal of A. Then the residue class algebra A/J is isometrically isomorphic to the field of complex numbers in the sense that each $\varphi(x) \in A/J$ is uniquely represented as $\varphi(x) = \lambda\varphi(e)$ where λ is a complex number depending on $\varphi(x)$.

Proof. By the Corollary of Theorem III.3.4, A/J is a commutative Banach algebra with unit in which every non-zero element is invertible. The fact that A/J is isometrically isomorphic to the complex field now follows by the Gelfand-Mazur Theorem.

Theorem 6. If A is a commutative Banach algebra with unit e, and J is a maximal ideal of A, then J is the kernel of a homomorphism of A onto the complex field C.

Proof. Let φ denote the natural homomorphism of A onto the field A/J, which has kernel J. Let Ψ denote the isomorphism (of Theorem III. 3.5) of A/J onto the complex field C. The scheme is illustrated by the diagram:

$$A \xrightarrow[\text{onto}]{\varphi} A/J \xrightarrow[\text{onto}]{\Psi} C \quad .$$

The composition $\Phi = \Psi \cdot \varphi$ (φ followed by Ψ) maps A onto C, and is easily verified to be a homomorphism of A onto C with kernel J.

Theorem 7. If A is a commutative Banach algebra with unit e, and λ is a homomorphism of A onto the complex field C, then the kernel of λ

is a maximal ideal J of A, and λ has the form $\Psi \cdot \varphi$ where φ is the natural homomorphism of A onto A/J, and Ψ is an isomorphism of A/J onto C.

Proof. Let J be the kernel of λ. Note that

$$J = \{x : x \in A, \lambda(x) = 0\} \quad .$$

It is easily verified that J is a proper ideal of A. (For example, if $x \in J$, $y \in A$, $\lambda(xy) = \lambda(x)\lambda(y) = 0$. Thus $xy \in J$.) That J is proper follows from the observation that λ maps A onto C. Suppose that J is not maximal. Then there exists a maximal ideal P of A which contains J. Consider the image $\lambda(P)$ of P. If $z_1, z_2 \in \lambda(P)$ and $\alpha, \beta \in C$, there exist elements $x_1, x_2 \in P$ such that $\lambda(x_1) = z_1$, $\lambda(x_2) = z_2$. But $\alpha\lambda(x_1) + \beta\lambda(x_2) = \lambda(\alpha x_1) + \lambda(\beta x_2) = \lambda(\alpha x_1 + \beta x_2)$. Since $\alpha x_1 + \beta x_2 \in P$, it follows that $\alpha\lambda(x_1) + \beta\lambda(x_2) = \alpha z_1 + \beta z_2 \in \lambda(P)$. Similarly, if $z \in \lambda(P)$ and $w \in C$, $wz \in \lambda(P)$. Hence $\lambda(P)$ is an ideal of C. We now show that $\lambda(P)$ is a proper ideal of C. Suppose $\lambda(P) = C$. Let $x \in A - P$. There exists on element $x_o \in P$ such that $\lambda(x_o) = \lambda(x)$. Thus $\lambda(x - x_o) = 0$, which implies that $x - x_o \in J \subset P$. This implies that $x \in P$, since $x = (x - x_o) + x_o$, and contradicts $x \in A - P$. It follows that $\lambda(P)$ is a proper ideal of C, properly containing the ideal $\{0\}$ of C. But this contradicts the fact that $\{0\}$ is the only proper ideal of C. Hence J is a maximal ideal. If φ is the natural homomorphism of A onto A/J, define $\Psi : A/J \to C$ by $\Psi(x + J) = \lambda(x)$. It is easily verified that Ψ is well defined and that Ψ is a homomorphism of A/J onto C. If $\lambda(x) = \lambda(y)$, then $\lambda(x - y) = 0$, and $x - y \in J$, which implies that $x + J = y + J$. Hence Ψ is an isomorphism of A/J onto C. Thus we have $\lambda = \Psi \cdot \varphi$.

The maximal ideals of A have been characterized in the two preceding theorems by the set of all homomorphisms Φ of A onto C, that is, the set of all multiplicative linear functionals on A which are not identically zero.

Definition 6. Let M denote the set of all maximal ideals of the commutative Banach algebra A with unit e. M will be called the maximal ideal space of A. Let x be a fixed element of A, and define the mapping $\hat{x} : M \rightarrow C$ by

$$\hat{x}(J) = \Phi[J](x) \text{ for every } J \in M.$$

Here $\Phi[J]$ is the homomorphism Φ of Theorem 6, the notation now altered to denote dependence of Φ on J. The mapping $\hat{x}$ will be called the representative function of the element $x \in A$. The mapping $x \rightarrow \hat{x}$ will be denoted by Δ, and is called the Gelfand representation of the commutative Banach algebra A with unit e.

In the following theorem we give some fundamental properties of the representative functions $\hat{x}$.

Theorem 8. Let A be a commutative Banach algebra with unit e. Let $x_1, x_2 \in A$, let $J_1, J_2 \in M$, and let $\lambda \in C$:

$$\text{If } x = x_1 + x_2, \text{ then } \hat{x}(J) = \hat{x}_1(J) + \hat{x}_2(J) \quad (J \in M). \tag{13}$$

$$\text{If } x = \lambda x_1, \text{ then } \hat{x}(J) = \lambda\hat{x}_1(J) \quad (J \in M). \tag{14}$$

If $x = x_1 x_2$, then $\hat{x}(J) = \hat{x}_1(J)\hat{x}_2(J) \quad (J \in M)$. (15)

$\hat{e}(J) = 1$ for every $J \in M$. (16)

$\hat{x}(J) = 0$ if and only if $x \in J$. (17)

If $J_1 \neq J_2$, then there exists an element $x \in A$ such that $\hat{x}(J_1) \neq \hat{x}(J_2)$. (18)

$|\hat{x}(J)| \leq \|x\|$ on M for every $x \in A$. (19)

Proof. For each fixed $J \in M$, $\hat{x}(J) = \Phi[J](x)$, and the mapping $\Phi[J] : A \to C$ is a homomorphism. The image of $e \in A$ under $\Phi[J]$ is the unit 1 of the complex field. The statements (13), (14), (15) and (16) follow directly from the above observations. For example, for (15),

$$\hat{x}(J) = \Phi[J](x_1 x_2) = \Phi[J](x_1) \cdot \Phi[J](x_2) = \hat{x}_1(J) \cdot \hat{x}_2(J) .$$

Since $\Phi[J]$ has J as its kernel, an element $x \in A$ maps into 0 under this homomorphism if and only if $x \in J$; this yields property (17). To show (18), suppose that $J_1, J_2 \in M$, and $J_1 \neq J_2$. Then there exists an $x \in A$ with $x \in J_1$ and $x \notin J_2$. Consequently, $\hat{x}(J_1) = 0$ and $\hat{x}(J_2) \neq 0$, which means that there is an element $x \in A$ for which $\hat{x}(J_1) \neq \hat{x}(J_2)$. The same type of argument applies in the case $x \in J_2$ and $x \notin J_1$. Finally, note that

$$|\hat{x}(J)| = |\Phi[J](x)| = |(\Psi\cdot\varphi)[J](x)| = \|\varphi[J](x)\| ,$$

since $\Psi[J]$ is an isometry, by the Gelfand-Mazur theorem. Therefore $|\hat{x}(J)| = \inf\{\|y\| : y \in \varphi[J]\}$, by definition of the quotient norm of the residue class $\varphi(x) = \varphi[J](x) \in A/J$. This proves the assertion (19).

Theorem 9. An element $x \in A$ is invertible if and only if the representative function $\hat{x} : M \to C$ vanishes nowhere on M.

Proof. By the Corollary of Theorem III.1.1, the element $x \in A$ has an inverse if and only if x is contained in no maximal ideal of A. This means that, for each maximal ideal J, $\hat{x}(J) \neq 0$, by (17).

Theorem 10. For each fixed $x \in A$, the range of the representative function $\hat{x}$ is precisely the spectrum of x.

Proof. Let $\lambda_o \in \sigma(x)$. Then $x - \lambda_o e$ is not invertible. Therefore, by Theorem III.3.9, $\hat{x} - \lambda_o \hat{e}$ must vanish on some maximal ideal J_o. Thus $(\hat{x} - \lambda_o \hat{e})(J_o) = 0 = \hat{x}(J_o) - \lambda_o \hat{e}(J_o) = \hat{x}(J_o) - \lambda_o$. Therefore $x(J_o) = \lambda_o$, and λ_o is an element of range of $\hat{x}$. Conversely, suppose that there exists a $J_o \in M$ such that $\hat{x}(J_o) = \lambda_o$, that is, λ_o is in the range of $\hat{x}$. Then $(\hat{x} - \lambda_o \hat{e})(J_o) = 0$, and therefore $x - \lambda_o e$ is not invertible (again by Theorem III.3.9). This means that $\lambda_o \in \sigma(x)$.

Remark 2. Since the range of the representative function $\hat{x}$ coincides with the spectrum of $x \in A$, the spectral radius of x is also given by

$$\rho(x) = \sup\{|\hat{x}(J)| : J \in M\} .$$

Theorem 11 (Spectral Radius Formula). Let A be a commutative Banach algebra with unit e. For every $x \in A$, $\lim_{n \to \infty} \|x^n\|^{1/n}$ exists, and

$$\rho(x) = \lim_{n \to \infty} \|x^n\|^{1/n} .$$

Proof. Put $\alpha = \sup\{|\hat{x}(J)| : J \in M\} = \rho(x)$, by (4). It will be shown that

$$\limsup_{n \to \infty} \|x^n\|^{1/n} \leq \alpha \leq \liminf_{n \to \infty} \|x^n\|^{1/n} , \tag{20}$$

which clearly will force $\liminf_{n \to \infty} \|x^n\|^{1/n} = \limsup_{n \to \infty} \|x^n\|^{1/n}$, and hence the conclusion. By (19) it follows that

$$\|x^n\| \geq \sup\{|\hat{x}^n(J)| : J \in M\} = \sup\{|\hat{x}(J)|^n : J \in M\} = \alpha^n .$$

Hence $\|x^n\|^{1/n} \geq \alpha$ for every positive integer n, which implies the right-hand inequality in (20). Now note that by Definition III.1.2 and Theorem III.3.10 if $\mu \in C$ and $|\mu| > \alpha$, then $x - \mu e$ is an invertible element of A; that is, $\mu \in r(x)$. Hence the function F defined by

$$F(\lambda) = (e - \lambda x)^{-1} = -(1/\lambda)[x - (1/\lambda)e]^{-1} = \lambda^{-1}R_x(\lambda^{-1}) \tag{21}$$

is strongly holomorphic for $|\lambda| < 1/\alpha$. By the Neumann series (see Lemma III.3.1)

$$(e - \lambda x)^{-1} = \sum_{n=0}^{\infty} (-1)^n (-\lambda x)^n = \sum_{n=0}^{\infty} \lambda^n x^n , \qquad (22)$$

which is valid for $\|\lambda x\| < 1$. Hence the Taylor series for F must be identical with (22):

$$F(\lambda) = \sum_{n=0}^{\infty} \lambda^n x^n, \ \|\lambda n\| < 1 . \qquad (23)$$

By convergence of the above Taylor series (23), it follows that

$$\lim_{n \to \infty} \|\lambda^n x^n\| = 0 \ \text{ if } |\lambda| < 1/\alpha .$$

Therefore, for suitably large positive integers n,

$$\|x^n\| = |\lambda|^{-n} \|\lambda^n x^n\| < |\lambda|^{-n} \text{ when } |\lambda| < 1/\alpha .$$

This means that $\limsup_{n \to \infty} \|x^n\|^{1/n} \leq |\lambda|^{-1}$, for every complex number λ with $\alpha < |\lambda|^{-1}$; that is,

$$\limsup_{n \to \infty} \|x^n\|^{1/n} \leq \alpha . \qquad (24)$$

This is the left-hand inequality in (20), and the proof is completed.

<u>Definition 7</u>. An element x of the commutative Banach algebra A

with unit e is said to be a <u>generalized nilpotent element</u> of A if

$$\lim_{n \to \infty} \|x^n\|^{1/n} = 0 \ .$$

<u>Corollary 4</u>. The radical $R = \bigcap_{J \in M} J$ of A coincides with the set of all generalized nilpotent elements of A.

This corollary is a direct consequence of the Spectral Radius Formula (and Remark 2) whereby

$$\sup\{|\hat{x}(J)| \ : J \in M\} = \lim_{n \to \infty} \|x^n\|^{1/n} \ .$$

4. The Structure Space of a Commutative Banach Algebra

Throughout the following discussion we shall denote by $\Delta(A)$ the totality of representative functions $\hat{x}$ which correspond by the Gelfand representation Δ to the elements x of the commutative Banach algebra A with unit e. Recall that each representative function $\hat{x} \in \Delta(A)$ is a mapping of M, the set of all maximal ideals of A, into the complex field C. The family of functions $\Delta(A)$ can be used to induce a topology in the set M as follows.

<u>Definition 1</u>. A <u>neighborhood</u> of an element $J_o \in M$ will be defined as any subset of M of the form

$$U(J_o; x_1, x_2, \ldots, x_n; \varepsilon)$$

$$= \{J : J \in M, |\hat{x}_i(J) - \hat{x}_i(J_o)| < \epsilon, i = 1, 2, \ldots, n\} ,$$

where the positive integer n, the positive real number ϵ, and the representative functions $\hat{x}_i \in \Delta(A)$ are arbitrary. The set of all such neighborhoods is a basis for a topology in M (often called the <u>Gelfand topology</u> in M) called the <u>weak topology</u> defined by the family of functions $\Delta(A)$. The topological space M thus obtained is called the <u>structure space</u> of the commutative Banach algebra A with unit e.

Note that M is then a Hausdorff space. This is established in Remark 2, by use of the result (18) of Theorem III.3.8, according to which the functions $\hat{x}$ <u>separate</u> points of M. The terminology in Definition 1 is explained in the following remarks, which apply also to more general situations.

<u>Remark 1</u>. If X is a space (nonempty set), and $\{f_\alpha : \alpha \in \Gamma\}$ is a nonempty family of complex-valued functions on X, there is a weakest topology τ_o in X for which all the functions f_α are continuous. There is at least one topology in X for which all the f_α are continuous, since the discrete topology (in which <u>all</u> subsets of X are open) satisfies the requirement trivially. The weakest topology in X for which the f_α are continuous is the intersection of all topologies τ in X for which the f_α are continuous. It is possible to describe τ_o, called the <u>weak topology generated by the functions</u> f_α, in a more explicit way. If f_α is to be continuous on X, it is necessary and sufficient that the inverse images $f_\alpha^{-1}(D) \in \tau_o$ for every open disk D in the complex plane. Consequently τ_o must contain all subsets of X of the form

$$\{x : x \in X, |f_\alpha(x) - f_\alpha(x_o)| < \varepsilon\} \quad ,$$

where $x_o \in X$, $\alpha \in \Gamma$, and $\varepsilon > 0$. Thus τ_o (which is closed under finite intersections) must contain all subsets of X of the form

$$V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon)$$

$$= \{x: x \in X, |f_{\alpha_1}(x) - f_{\alpha_1}(x_o)| < \varepsilon, \ldots, |f_{\alpha_n}(x) - f_{\alpha_n}(x_o)| < \varepsilon\},$$

where $x_o \in X$, $\varepsilon > 0$, and $\{\alpha_1, \ldots, \alpha_n\}$ is a finite subset of the indexing set Γ. Note that $x_o \in V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon)$. Given the sets $V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon)$ and $V(x_o; f_{\beta_1}, \ldots, f_{\beta_m}; \delta)$, the set

$$V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}, f_{\beta_1}, \ldots, f_{\beta_m}; \varepsilon') \subset V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon)$$

$$\cap\, V(x_o; f_{\beta_1}, \ldots, f_{\beta_m}; \delta) \text{ if } \varepsilon' = \min\{\varepsilon, \delta\} \quad .$$

Furthermore, if $x_1 \in V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon)$, and

$$\varepsilon_1 = \varepsilon - \max\{|f_{\alpha_1}(x_1) - f_{\alpha_1}(x_o)|, \ldots, |f_{\alpha_n}(x_1) - f_{\alpha_n}(x_o)|\} \quad ,$$

then

$$V(x_1; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon_1) \subset V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \varepsilon) \quad .$$

To deduce this inclusion, observe that, if $x \in V(x_1; f_{\alpha_1}, \ldots, f_{\alpha_n}; \epsilon_1)$, then for each $i = 1, \ldots, n$,

$$|f_{\alpha_i}(x) - f_{\alpha_i}(x_o)| \le |f_{\alpha_i}(x) - f_{\alpha_i}(x_1)| + |f_{\alpha_i}(x_1) - f_{\alpha_i}(x_o)|$$

$$< \epsilon_1 + \max\{|f_{\alpha_i}(x_1) - f_{\alpha_i}(x_o)| : i = 1, \ldots, n\} = \epsilon,$$

by the definition of ϵ_1. Thus $x \in V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \epsilon)$, and the inclusion is verified. Thus the sets of the form $V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \epsilon)$ form a neighborhood base for τ_o at the point x_o. The weak topology τ_o generated by the functions f_α is composed of arbitrary unions of sets of the form $V(x_o; f_{\alpha_1}, \ldots, f_{\alpha_n}; \epsilon)$ where $x_o \in X$, $\{\alpha_1, \ldots, \alpha_n\}$ is a finite subset of Γ, and $\epsilon > 0$.

Remark 2. It should be pointed out that X with the weak topology τ_o is a Hausdorff space if, whenever $x,y \in X$ and $x \ne y$, there exists an $\alpha \in \Gamma$ (α depending on x,y) such that $f_\alpha(x) \ne f_\alpha(y)$. In this case the functions f_α are said to separate points of X. To deduce this, let $\epsilon = \frac{1}{2}|f_\alpha(x) - f_\alpha(y)|$, and consider the neighborhoods $V(x; f_\alpha; \epsilon)$, $V(y; f_\alpha; \epsilon)$. If $z \in V(x; f_\alpha; \epsilon)$, then it follows from the inequality

$$2\epsilon = |f_\alpha(y) - f_\alpha(x)| \le |f_\alpha(y) - f_\alpha(z)| + |f_\alpha(z) - f_\alpha(x)|$$

$$< |f_\alpha(y) - f_\alpha(z)| + \epsilon$$

that $|f_\alpha(z) - f_\alpha(y)| > \epsilon$, which implies that $z \notin V(y; f_\alpha; \epsilon)$. Thus

$V(x; f_\alpha; \epsilon)$, $V(y; f_\alpha; \epsilon)$ are disjoint neighborhoods.

Remark 3. The concept of the weak topology generated by a family of functions may also be considered in situations more general than that considered above. If X_α is a topological space for each α in a nonempty indexing set Γ, the Cartesian product space $X = \bigtimes_{\alpha\in\Gamma} X_\alpha$ is, by definition, the set of all functions x with domain Γ such that $x(\alpha) \in X_\alpha$ for each $\alpha \in \Gamma$. Often $x(\alpha)$ is written in the form x_α, and x_α is called the α^{th} coordinate of x. For each $\alpha \in \Gamma$, there is a function p_α on X defined by

$$p_\alpha(x) = x_\alpha .$$

The function p_α is called the projection of X onto X_α. The product topology on X is the weak topology generated by the family of projections p_α, and is the weakest topology for which all the projections are continuous functions. A very important theorem in general topology is the theorem of Tychonoff which asserts that if $\{X_\alpha : \alpha \in \Gamma\}$ is a nonempty collection of compact topological spaces and the Cartesian product space $X = \bigtimes_{\alpha\in\Gamma} X_\alpha$ is given the product topology, then X is compact. For a proof of this theorem, reference may be made to Kelley [13], p. 143 or Royden [24], pp. 144-145.

It is helpful at this point to recall some basic facts about Banach spaces. If X is a Banach space, the dual space X^* of X is the space of all bounded complex-valued linear functionals on X with norm defined by

$$\|F\| = \sup\{|F(x)| : \|x\| \le 1\} \quad (F \in X^*) .$$

With this norm, the space X^* is also a Banach space. Its dual space $(X^*)^*$ is called the <u>second dual space of X</u>, and is usually denoted X^{**}.

For each $x \in X$, let Λ_x be the functional defined on the domain X^* by $\Lambda_x(F) = F(x)$. It is easily verified that Λ_x is linear on X^* for each $x \in X$. If $\|F\| \le 1$,

$$|\Lambda_x(F)| = |F(x)| \le \|F\| \cdot \|x\| \le \|x\| .$$

Hence Λ_x is, for each $x \in X$, a bounded linear functional on X^*, and $\|\Lambda_x\| \le \|x\|$. Thus $\Lambda_x \in X^{**}$ for each $x \in X$. An argument based on the Hahn-Banach theorem shows that $\|\Lambda_x\| = \|x\|$. The mapping $x \to \Lambda_x$ is a linear space homomorphism of X into X^{**}, since $\Lambda_{x+y} = \Lambda_x + \Lambda_y$ and $\Lambda_{\alpha x} = \alpha\Lambda_x$ for all $\alpha \in C$, and $x,y \in X$. This mapping is called the <u>natural mapping</u> of X into X^{**}.

The sets of the form

$$W(x_o; F_1, \ldots, F_n; \varepsilon) = \{x : x \in X, |F_i(x) - F_i(x_o)| < \varepsilon \text{ for } i = 1,\ldots,n\} ,$$

where $x_o \in X$, $F_1, \ldots, F_n \in X^*$ and $\varepsilon > 0$ form a neighborhood base at x_o for each $x_o \in X$. This is shown just as in Remark 1 (p. 54). The associated topology in X is the weak topology generated by the functions $F \in X^*$, and X with this topology is a Hausdorff space. One can define, similarly, the weak topology in X^* generated by the functions in X^{**}.

But another topology on X^* has, it turns out, greater importance. This topology, called the <u>weak* topology in X^*</u> is the weak topology generated by the subset of X^{**} which is the image of X under the natural mapping $x \to \Lambda_x$. (For non-reflexive Banach spaces, the natural mapping carries X onto a proper subset of X^{**}.). For the weak* topology a typical neighborhood at F_o has the form

$$W(F_o; \Lambda_{x_1}, \ldots, \Lambda_{x_n}; \epsilon) = \{F : F \in X^*, |\Lambda_{x_i}(F) - \Lambda_{x_i}(F_o)| < \epsilon$$
$$\text{for } i = 1, \ldots, n\} \quad .$$

Using the definition of Λ_{x_1} it follows that

$$W(F_o; \Lambda_{x_1}, \ldots, \Lambda_{x_n}; \epsilon) = \{F : F \in X^*, |F(x_i) - F_o(x_i)| < \epsilon$$
$$\text{for } i = 1, \ldots, n\} \quad . \tag{1}$$

The space X^* with the weak* topology is a Hausdorff space. For more details regarding the above remarks, reference may be made to Lorch [15].

One of the most important theorems concerning the weak* topology in X^* is the following theorem of Alaoglu.

<u>Theorem 1</u>. (<u>Alaoglu's Theorem</u>). The closed unit ball $S^* = \{F : F \in X^*, \|F\| \leq 1\}$ in X^* is compact in the weak* topology.

<u>Proof</u>. For a fixed $x \in X$, and any $F \in S^*$,

$$|F(x)| \leq \|F\| \cdot \|x\| \leq \|x\| \quad . \tag{2}$$

The values assumed by the functionals F at x are contained in the closed disk $K_x = \{z : |z| \leq \|x\|\}$ in the complex plane. By the Tychonoff theorem, as stated in Remark 3 (p. 57), the product space $P = \times_{x \in X} K_x$ is compact in the product topology on P. The points of P are the functions h with domain X such that $h(x) \in K_x$ for each $x \in X$. Thus, on account of (2), $S^* \subset P$. A typical neighborhood at $h_o \in P$ in the product topology is given by

$$W(h_o; p_{x_1}, p_{x_2}, \ldots, p_{x_n}; \varepsilon) = \{h : h \in P, |p_{x_i}(h) - p_{x_i}(h_o)| < \varepsilon$$

$$\text{for } i = 1, \ldots, n\}$$

$$= \{h : h \in P, |h(x_i) - h_o(x_i)| < \varepsilon \text{ for } i = 1, \ldots, n\},$$

since $p_x(h) = h(x)$, the x^{th} coordinate of h. The set S^*, viewed as a subset of P, has an induced topology, induced by the product topology on P. A typical basis set in this <u>induced topology</u> has the form

$$W(h_o; p_{x_1}, \ldots, p_{x_n}; \varepsilon) \cap S^*$$

$$= \{h : h \in S^*, |h(x_i) - h_o(x_i)| < \varepsilon \text{ for } i = 1, \ldots, n\} , \tag{3}$$

where $h_o \in S^*$. Upon comparing (1) and (3), it is seen that the topology induced on S^* by the product topology in P is the same as the topology induced on S^* by the weak* topology in X^*. The subsets of S^* specified as open are the same in each case.

Since, by a standard theorem in general topology, a closed subset of a compact Hausdorff space is compact, it suffices to prove that S^* is a closed subset of P. Suppose $g_o \in P$ is a limit point of S^*, and let $\epsilon > 0$ be given. Every open set in P containing g_o then contains points of S^*. Suppose that $x, y \in X$. The neighborhood $W(g_o; p_x, p_y, p_{x+y}; \epsilon)$ of g_o contains points of S^*. Let $g \in S^*$ be such a point. Since g belongs to X^*, g is a bounded linear functional on X, and hence $g(x + y) = g(x) + g(y)$. Since $g \in W(g_o; p_x, p_y, p_{x+y}; \epsilon)$ and $p_x(g) = g(x)$, it follows that

$$|g(x) - g_o(x)| < \epsilon, \quad |g(y) - g_o(y)| < \epsilon,$$

and

$$|g(x + y) - g_o(x + y)| < \epsilon \quad .$$

Thus

$$|g_o(x + y) - g_o(x) - g_o(y)|$$

$$= |g_o(x + y) - g(x + y) + g(x) + g(y) - g_o(x) - g_o(y)|$$

$$\leq |g_o(x + y) - g(x + y)| + |g(x) - g_o(x)| + |g(y) - g_o(y)| < 3\varepsilon \quad .$$

Since g_o does not depend on ε and $\varepsilon > 0$ is arbitrary, it follows that

$$g_o(x + y) = g_o(x) + g_o(y), \text{ for all } x,y \in X \quad .$$

A similar argument shows that $g_o(\alpha x) = \alpha g_o(x)$ for $\alpha \in C$, and $x \in X$. Thus g_o is a linear functional on X. For each $x \in X$ and $\varepsilon > 0$, there exists $g \in S^*$ such that $g \in W(g_o; p_x; \varepsilon)$, that is, such that $|g(x) - g_o(x)| < \varepsilon$. Hence

$$|g_o(x)| \leq |g_o(x) - g(x)| + |g(x)| < \varepsilon + \|g\| \cdot \|x\| \leq \varepsilon + \|x\| \quad .$$

If $\|x\| \leq 1$, $|g_o(x)| < \varepsilon + 1$. Thus $\|g_o\| \leq \varepsilon + 1$, for every $\varepsilon > 0$. Hence $\|g_o\| \leq 1$. Consequently $g_o \in S^*$. Thus S^* is a closed subset of P. The proof is complete.

<u>Note</u>. The above proof of the Alaoglu theorem is a much expanded version of that in Loomis [14], p. 22, or that in Lorch [15], pp. 26-27.

<u>Theorem 2</u>. The structure space (maximal ideal space with the Gelfand topology) M of a commutative Banach algebra A with unit e is a compact Hausdorff space. For each $x \in A$, the function $\hat{x} : M \to C$ is continuous.

<u>Proof</u>. By Theorem III.3.8, the functions $\hat{x}$ separate points of M. Thus M is a Hausdorff space. The weak topology generated by the functions $\hat{x}$ for $x \in A$ is specified by prescribing a neighborhood base at each

$J_o \in M$, with typical neighborhood

$$U(J_o; \hat{x}_1, \ldots, \hat{x}_n; \varepsilon) = \{J : J \in M, |\hat{x}_i(J) - \hat{x}_i(J_o)| < \varepsilon \text{ for } i = 1, \ldots, n\},$$

where $x_1, \ldots, x_n \in A$ and $\varepsilon > 0$. This topology in M is, it should be recalled, the weakest topology in M for which all the functions $\hat{x}$ for $x \in A$ are continuous. Viewed as a Banach space, A has the dual space A^*. For each fixed $J \in M$, let F_J: $A \to C$ be defined by

$$F_J(x) = \hat{x}(J) \quad .$$

Using the known properties of $\hat{x}$, it is now shown that each functional F_J: $A \to C$ is a bounded linear functional on A and, consequently, $F_J \in A^*$, for each $J \in M$. If $\alpha, \beta \in C$ and $x, y \in A$,

$$F_J(\alpha x + \beta y) = \widehat{\alpha x + \beta y}(J) = \alpha\hat{x}(J) + \beta\hat{y}(J) = \alpha F_J(x) + \beta F_J(y) \quad ,$$

by conditions III.3(13), 3(14). Then F_J is linear. If $\|x\| \le 1$, it follows from condition III.3(19) that

$$|F_J(x)| = |\hat{x}(J)| \le \|x\| \le 1.$$

Hence F_J is a bounded linear functional on A, with $\|F_J\| \le 1$, in fact. This holds for each $J \in M$. Thus the set

$$K = \{F_J : J \in M\} \subset S^* \quad ,$$

where S^* is the closed unit sphere $\{F : F \in A^*, \|F\| \le 1\}$ in A^*. By the theorem of Alaoglu S^* is compact in the weak* topology. It is now shown that K is closed in the weak* topology in A^*. Note first that, in any case, $\overline{K} \subset \overline{S^*} = S^*$. Suppose $F_o \in \overline{K}$. It must be shown that $F_o \in K$. For each $\varepsilon > 0$, the neighborhood (in S^*)

$$\{F : F \in S^*, |F(e) - F_o(e)| < \varepsilon \}$$

contains a point of K. Thus, for some $J \in M$, $|F_J(e) - F_o(e)| < \varepsilon$. But $F_J(e) = \hat{e}(J) = 1$ for every $J \in M$. Thus $|1 - F_o(e)| < \varepsilon$ for every $\varepsilon > 0$, and consequently $F_o(e) = 1$. Next, we show that $F_o(xy) = F_o(x)F_o(y)$ for all $x, y \in A$. Let $x, y \in A$ and suppose $\varepsilon > 0$. The neighborhood of F_o in S^* defined by

$$W(F_o; \Lambda_x, \Lambda_y, \Lambda_{xy}; \varepsilon) \cap S^*$$

$$= \{F : F \in S^*, |F(x)-F_o(x)|<\varepsilon, |F(y)-F_o(y)|<\varepsilon, |F(xy)-F_o(xy)|<\varepsilon\}$$

contains an element F_J of K. Thus $|F_J(x) - F_o(x)| < \varepsilon$, $|F_J(y) - F_o(y)| < \varepsilon$, and $|F_J(xy) - F_o(xy)| < \varepsilon$. But $F_J(xy) = \widehat{xy}(J) = \hat{x}(J)\hat{y}(J)$, by condition III.3(15). Now

$$|F_o(xy) - F_o(x)F_o(y)|$$

$$= |F_o(xy) - F_J(xy) + F_J(x)F_J(y) - F_J(x)F_o(y) + F_J(x)F_o(y) - F_o(x)F_o(y)|$$

$$\leq |F_o(xy) - F_J(xy)| + |F_J(x)|\ |F_J(y) - F_o(y)| + |F_o(y)|\ |F_J(x) - F_o(x)| .$$

Since $|F_J(x)| = |\hat{x}(J)| \leq \|x\|$ by condition III.3(19), and $|F_o(y)| \leq \|F_o\| \cdot \|y\|$, it follows that

$$|F_o(xy) - F_o(x)F_o(y)| \leq \varepsilon(1 + \|x\| + \|y\|) \quad .$$

Since this is true for each $\varepsilon > 0$, then $F_o(xy) = F_o(x)F_o(y)$ and F_o is multiplicative, with $F_o(e) = 1$. Hence there exists a maximal ideal $J_o \in M$, such that $F_o = F_{J_o}$, and it has been shown that $F_o \in K$. Hence K is a closed subset of S^* in the weak* topology. Thus K is compact in the weak* topology, and K, regarded as a topological space, is compact in the weak* topology. A typical neighborhood of F_J in K with this topology is

$$\{F_J : J \in M, |F_J(x_i) - F_{J_o}(x_i)| < \varepsilon \text{ for } i = 1, \ldots, n\}$$

$$= \{F_J : J \in M, |\hat{x}_i(J) - \hat{x}_i(J_o)| < \varepsilon \text{ for } i = 1, \ldots, n\} \quad .$$

Referring to the neighborhoods used to define the Gelfand topology in the maximal ideal space M, it follows upon identifying J with F_J that the open sets in M are precisely those induced by the open sets in K with the weak* topology. An open cover of M induces an open cover of K. Since K has been shown to be compact, the open cover induced on K has a finite subcover which in turn produces a corresponding subcover of M. Consequently M with the Gelfand topology is a compact Hausdorff space.

Note. The above proof follows the general lines of proofs in Loomis [14], p. 52 and Bachman and Narici [1], p. 341, but is in a considerably expanded form. A somewhat different proof appears in Naimark [19], p. 197, not using the Alaoglu theorem explicitly but essentially involving the same ideas, including the fundamental use of the Tychonoff theorem on product topologies.

The fundamental properties of the Gelfand representation Δ of the commutative Banach algebra A with unit e can now be summarized in the following theorem.

Theorem 3. A commutative Banach algebra A with unit e is represented homomorphically by the algebra $\Delta(A)$ of continuous functions $\hat{x}$ on the compact Hausdorff space M of all maximal ideals J of A with the Gelfand topology. The radical R of A consists of precisely those elements which are represented in $\Delta(A)$ by functions identically zero on M. The representation $x \to \hat{x}$ is isomorphic if and only if A is semi-simple.

Proof. The Gelfand representation Δ is a homomorphism by Theorem III.3.8. Assume that $x \in R = \bigcap_{J \in M} J$. Then $x \in J$ for every $J \in M$. Hence $\hat{x}(J) = 0$ on M by condition III.3.(17). On the other hand, every function $\hat{x}$ which is identically zero represents an element $x \in R$. Thus, the radical R is the kernel of the homomorphism Δ. Since a homomorphism is an isomorphism if and only if its kernel is trivial, it follows that the

Gelfand representation is isomorphic if and only if the radical of A is the trivial ideal $\{0\}$.

Example 1. Let $C[0,1]$ denote the space of all complex-valued continuous functions on $[0,1]$ with the usual operations of scalar multiplication, addition, and multiplication. Define the norm of an element x of $C[0,1]$ by

$$\|x\| = \sup\{|x(t)| : 0 \le t \le 1\} \quad . \tag{4}$$

Then $C[0,1]$ is a commutative Banach algebra with unit. Let t_o be a fixed point in $[0,1]$, and define

$$J_{t_o} = \{x \in C[0,1] : x(t_o) = 0\} \quad . \tag{5}$$

It will be shown that : (i) J_{t_o} is a maximal ideal of $C[0,1]$, and (ii) every maximal ideal in $C[0,1]$ is of the form (5). Note that J_{t_o} is a linear subspace, and that $x \in J_{t_o}$ implies that $x(t_o)y(t_o) = 0$ for every $y \in C[0,1]$; therefore J_{t_o} is an ideal. To show that J_{t_o} is maximal, define the homomorphism

$$\Phi_{t_o} : C[0,1] \to C \text{ by}$$

$$\Phi_{t_o}(x) = x(t_o) \text{ for every } x \in C[0,1] \quad . \tag{6}$$

This homomorphism has J_{t_o} as its kernel and hence, as in Theorem III.3.7,

J_{t_o} is maximal. To prove the second assertion (ii), suppose that J_o is any maximal ideal of $C[0,1]$. It will be shown that there exists a point t_o in $[0,1]$ for which $J_o = J_{t_o}$. Assume that the contrary holds; that is, for every point τ in $[0,1]$ there exists an element x_τ in J_o such that $x_\tau(\tau) \neq 0$. Then corresponding to each τ in $[0,1]$ there is a neighborhood of τ, say $N(\tau)$, such that $x_\tau(t) \neq 0$ for every t in $N(\tau)$. This follows by continuity of x_τ. Since $[0,1]$ is a compact topological space, and since the neighborhoods $N(\tau)$ ($\tau \in [0,1]$) cover $[0,1]$, there is a finite number of neighborhoods $N(\tau_1)$, $N(\tau_2)$, ..., $N(\tau_n)$ which also cover $[0,1]$. Let $x_{\tau_1}, x_{\tau_2}, \ldots, x_{\tau_n}$ in J_o be the functions corresponding to the points $\tau_1, \tau_2, \ldots, \tau_n$. Then the function x defined by

$$x(t) = \sum_{i=1}^{n} x_{\tau_i}(t)\, \overline{x_{\tau_i}(t)} = \sum_{i=1}^{n} |x_{\tau_i}(t)|^2 \tag{7}$$

for every t in $[0,1]$, is also an element of J_o. But x does not vanish anywhere on $[0,1]$. Therefore $y(t) = 1/x(t)$ is continuous on the unit interval, and is consequently in $C[0,1]$. This shows that x has the inverse y in $C[0,1]$ since $x(t)y(t) = y(t)x(t) \equiv 1$ in t. Hence, by Corollary III.1.1, x cannot belong to a maximal ideal. This is a contradiction. Hence there must be some point t_o in the unit interval such that every element of J_o vanishes at t_o. Therefore $J_o \subset J_{t_o}$, which proves (ii). Note that the correspondence

$$t \rightarrow J_t$$

between the set of points [0,1] and the set of all maximal ideals of C[0,1] is one-to-one. The representative functions $\hat{x}$ defined by the Gelfand representation of C[0,1] are given by

$$\hat{x}(J_{t_o}) = \Phi[J_{t_o}](x) \quad \text{(as in Theorem III.3.6)},$$

and in view of (6),

$$\hat{x}(J_{t_o}) = \Phi_{t_o}(x) = x(t_o) \quad . \tag{8}$$

CHAPTER IV

THE SPECTRAL RESOLUTION OF NORMAL OPERATORS ON HILBERT SPACE

1. The Commutative Banach Algebra Generated by a Normal Operator

Let A(H) denote the Banach algebra with unit of all operators on the Hilbert space H.

Definition 1. Let N be a set of normal operators in A(H) which satisfies the conditions:

$$S, T \in N \text{ implies } ST = TS \text{ (commutativity)} \tag{1}$$

$$T \in N \text{ implies } T^* \in N \text{ (closure under adjoint formation).} \tag{2}$$

Let N' denote the set of all operators in A(H) which commute with every operator in N, and let $B = N'' = (N')'$ denote the set of all operators in A(H) which commute with every operator in N'.

Theorem 1. Let N be as in Definition 1. Then $N'' = (N')'$ is a commutative Banach algebra with unit I (the identity operator on H). The algebra $B = N''$ consists of normal operators also.

Proof. It is easily verified that if $T_1, T_2 \in N''$, and $\alpha, \beta \in C$, then the product (composition) $T_1T_2 \in N''$ and $\alpha T_1 + \beta T_2 \in N''$. We show next that $N'' \subseteq N'''$, where N''' is defined as $(N'')'$. If $T \in N$, then $TS = ST$ for all $S \in N$ by (1), and thus $T \in N'$. Hence $N \subseteq N'$. If $T \in N''$, then $TS = ST$ for all $S \in N'$ by definition of N'' as $(N')'$. Since $N \subseteq N'$, it follows that, if $T \in N''$, then $TS = ST$ for all $S \in N$, and hence $T \in N'$. Thus it has been shown that $N'' \subseteq N'$. If $T \in N''$, then $TS = ST$ for all $S \in N'$. But $N'' \subseteq N'$. Hence, if $T \in N''$, then $TS = ST$ for all $S \in N''$, and consequently $T \in N'''$. Thus $N'' \subseteq N'''$. It follows that multiplication in N'' is commutative since if $S, T \in N''$ then $S \in N''$ and $T \in N'''$, and this implies that $TS = ST$.

It is next shown that N'' is closed under formation of adjoints. If $R \in N'$, then $RS = SR$ for all $S \in N$. Note that, by (2), $S^* \in N$ whenever $S \in N$. Thus, if $R \in N'$, $RS^* = S^*R$ for all $S \in N$, and consequently $SR^* = R^*S$ for all $S \in N$, which is equivalent to saying that $R^* \in N'$. Suppose now that $T \in N''$. Then $TS = ST$ for all $S \in N'$. Hence $TS^* = S^*T$ for all $S \in N'$. By taking adjoints, it follows that $ST^* = T^*S$ for all $S \in N'$, and thus $T^* \in N''$. We have shown that $T \in N''$ implies that $T^* \in N''$. Hence N'' is closed under adjoint formation. It is clear that $I \in N'$ and $I \in N''$, since $IS = SI$ for every $S \in A(H)$.

It should be recalled that for each bounded linear operator T on H the <u>norm</u> of T is defined by $\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\}$. This norm is used in the commutative algebra $B = N''$ to yield a normed algebra B. With this norm, B is complete as a normed linear spaace. If $\{T_n\}$ is a Cauchy sequence in N'', then in particular $\{T_n\}$ is a Cauchy sequence in A(H). It is then a consequence of basic theorems on bounded linear operators on H that there exists a $T \in A(H)$ for which $\|T_n - T\| \to 0$ as $n \to \infty$. The limit operator T is defined by

$$Tx = s - \lim_{n \to \infty} T_n x \text{ for each } x \in H.$$

If $\{T_n\}$ is a Cauchy sequence in $B = N''$, then $T_n S = ST_n$ for all $S \in N'$. We wish to show that the limit operator T actually belongs to N''. For each $x \in H$, $(T_n S - TS)x = (T_n - T)Sx$, and

$$\|(T_n - T)Sx\| \leq \|T_n - T\| \cdot \|S\| \cdot \|x\| \to 0 \text{ as } n \to \infty.$$

Thus $s - \lim_{n \to \infty} T_n S = TS$. Similarly, $s - \lim_{n \to \infty} ST_n = ST$. Thus $T_n S = ST_n$ for $n = 1, 2, 3, \ldots$ implies that $TS = ST$. This is true for all $S \in N'$. Thus $T \in N''$. Hence $B = N''$ is a commutative Banach algebra with unit I. Since $N \subseteq A(H)$ consists of normal operators, every element of $B = N''$ is a normal operator. Indeed, if $T \in B$, then $T^* \in B$ and (since B is commutative) $T^*T = TT^*$. Thus T is a normal operator.

Remark 1. If T is a fixed normal operator on H, the set $N(T) = \{T, T^*\}$ satisfies trivially the conditions (1) and (2) of Definition 1. Thus $B = N(T)''$ has the properties of the set B of Theorem 1, and

$$\{0, I, T, T^*\} \subseteq B \quad .$$

Definition 2. The commutative Banach algebra $B = N(T)''$ with unit, where T is a fixed normal operator on H and $N(T) = \{T, T^*\}$, will be called the Banach algebra generated by the normal operator T.

2. The Gelfand Representation of the Banach Algebra Generated by a Normal Operator

Theorem 1. (Kakutani-Krein). Let X be a compact Hausdorff space, and let C(X) be the set of all complex-valued continuous functions on X. Let a subset D of C(X) satisfy the conditions:

(a) If f, g $\in$ D and α, β are complex numbers, then $\alpha f + \beta g \in D$ and the product f g $\in$ D.

(b) The constant function $1(x) \equiv 1$ belongs to D.

(c) If $\{f_n\}$ is a sequence of functions in D with a uniform limit f_∞, then $f_\infty \in D$.

Then D = C(X) if and only if we have:

(d) D separates points of X, and

(e) If f $\in$ D, then $\overline{f} \in D$, where $\overline{f}$ denotes the complex conjugate of f.

A proof of this corollary of the Stone-Weierstrass theorem is given in Yosida [30], p. 10. The terminology separates points of X is used in the same manner as that given in Remark 4.2 of Chapter III.

We shall apply Theorem IV.2.1 to show that the Gelfand representation Δ maps the Banach algebra B = N(T)$''$ isomorphically onto the set of all continuous, complex-valued functions on the structure space M of B.

Theorem 2. Let B be the commutative Banach algebra with unit I generated by the fixed normal operator $T_o \in A(H)$. By the Gelfand representation $\Delta : T \to \hat{T}$, the Banach algebra B is represented isomorphically by the algebra C(M) of all the continuous, complex-valued functions on the compact structure space M of all the maximal ideals J of B in such a way that, for any $T \in B$:

$$\|T\| = \sup\{|\hat{T}(J)| : J \in M\} \text{ (the spectral radius of T)} \qquad (1)$$

$\hat{T} : M \to C$ is real-valued if and only if the operator T is Hermitian. (2)

$\hat{T}(J) \geq 0$ on M if and only if T is Hermitian and positive. (3)

Proof. Let $T \in B$. By Theorem III.3.11 $\lim_{n \to \infty} \|T_n\|^{1/n} < \infty$, and

$$\sup\{|\hat{T}(J)| : J \in M\} = \lim_{n \to \infty} \|T^n\|^{1/n} . \qquad (4)$$

We will show that

$$\|T\| = \lim_{n \to \infty} \|T^n\|^{1/n} . \qquad (5)$$

By normality of T, it follows that $S = T^*T = TT^*$ is Hermitian. Hence, by Theorem II.3.2,

$$\|S\| = \|T^*T\| = \sup\{|(T^*Tx, x)| : \|x\| \leq 1\}$$

$$= \sup\{|(Tx, Tx)| : \|x\| \leq 1\} = \sup\{\|Tx\|^2 : \|x\| \leq 1\} \quad . \qquad (6)$$

Since $\|Tx\| \leq \|T\|$ if $\|x\| \leq 1$, $\|Tx\|^2 \leq \|T\|^2$ if $\|x\| \leq 1$. Thus

$$\alpha = \sup\{\|Tx\|^2 : \|x\| \leq 1\} \leq \|T\|^2 \quad .$$

Since $\|Tx\|^2 \leq \alpha$ for all x such that $\|x\| \leq 1$ implies that $\|Tx\| \leq \sqrt{\alpha}$ for all x such that $\|x\| \leq 1$, it follows that $\|T\| \leq \sqrt{\alpha}$, and $\|T\|^2 \leq \alpha$. It has been observed that $\alpha \leq \|T\|^2$, and hence

$$\alpha = \sup\{\|Tx\|^2 : \|x\| \leq 1\} = \|T\|^2 \quad .$$

Hence (6) implies that $\|S\| = \|T^*T\| = \|T\|^2$.

Since $(T^2)^* = (TT)^* = T^*T^* = (T^*)^2$, we have

$$\|T^2\|^2 = \|(T^2)^* T^2\| = \|(T^*)^2 T^2\| \quad .$$

Again, by the normality of T, it follows that

$$\|T^2\|^2 = \|T^*T^*TT\| = \|T^*TT^*T\| = \|(T^*T)^2\| = \|S^2\| \quad .$$

Since S^2 is Hermitian, it follows (again by Theorem II.3.2) that

$$\|S\|^2 = \sup\{\|Sx\|^2 : \|x\| \le 1\} = \sup\{|(Sx, Sx)| : \|x\| \le 1\}$$

$$= \sup\{|(S^2x, x)| : \|x\| \le 1\} = \|S^2\| .$$

Therefore

$$\|T^2\|^2 = \|S^2\| = \|S\|^2 = (\|T\|^2)^2 ,$$

or, equivalently,

$$\|T^2\| = \|T\|^2 . \qquad (7)$$

We have also shown that T^2 is normal if T is normal. This fact together with (7) yields the result

$$\|T^{2^k}\| = \|T\|^{2^k} \text{ for } k = 1, 2, \ldots,$$

by induction on k. Hence $\|T\| = \lim_{k \to \infty} \|T^{2^k}\|^{\frac{1}{2^k}}$. Since $\lim_{n \to \infty} \|T^n\|^{1/n}$ <u>exists</u>, it follows that formula (5) is verified. Noting (4), we have proved assertion (1). Since $\|T\| = 0$ if and only if $T = \mathbf{0}$, it follows from Corollary III.3.4 that the algebra B is semi-simple. Thus the mapping $T \to \hat{T}$ is an <u>isomorphism</u> by Theorem III.4.3.

To show necessity in (2), suppose that $T \in B$ is Hermitian but $\hat{T}(J_o) = a + ib$, with a, b real and $b \ne 0$, for some $J_o \in M$. Put

$$S = \frac{T - aI}{b} \in B \quad ,$$

and note that S is also Hermitian. By Theorem III.3.8 we have

$$(\widehat{I + S^2})\ (J_o) = \hat{I}(J_o) + \widehat{S^2}(J_o) = \hat{I}(J_o) + (\hat{S}(J_o))^2$$

$$= 1 + \frac{(a + ib)^2}{b^2} - \frac{2a}{b^2}\ (a + ib) + \frac{a^2}{b^2} = 1 + i^2 = 0.$$

By Theorem III.3.9, $I + S^2$ is not invertible in B, since $\widehat{I + S^2}$ vanishes at $J_o \in M$. This contradicts Theorem II.3.4. This theorem asserts, in fact, that the operator $I + S^*S$ (which is the same as $I + S^2$, since S is Hermitian) is invertible in A(H). Since $P = I + S^2 \in B = N''$, PR = RP for all $R \in N'$. Thus $P^{-1}(PR)P^{-1} = P^{-1}(RP)P^{-1}$ for all $R \in N'$, and thus $RP^{-1} = P^{-1}R$ for all $R \in N'$. Consequently $P^{-1} = (I + S^2)^{-1} \in B$.

To show that if $\hat{T}$ is real-valued then T is Hermitian, suppose that $T \in B$ and T is not Hermitian. Write

$$T = \frac{T + T^*}{2} + i\,\frac{T - T^*}{2i} \quad ,$$

where the operators $\frac{1}{2}(T + T^*)$ and $\frac{1}{2i}(T - T^*)$ are the unique constituents of the Hermitian decomposition of T, as mentioned in Remark II.3.3. Since $\frac{1}{2}(T + T^*)$ is Hermitian, the Hermitian operator $\frac{1}{2i}(T - T^*)$ is not the zero operator **0**, since if that were the case T would be Hermitian, contrary to assumption. By the isomorphism $T \rightarrow \hat{T}$, it follows that there must exist an element $J_o \in M$ such that

$$\widehat{\left(\frac{T - T^*}{2i}\right)}(J_o) \neq 0 \quad .$$

Consequently

$$\hat{T}(J_o) = \frac{\hat{T} + \widehat{T^*}}{2}(J_o) + i\,\frac{\hat{T} - \widehat{T^*}}{2i}(J_o)$$

is not a real number. since by the earlier part of this argument, $\frac{\hat{T} + \widehat{T^*}}{2}(J_o)$ and $\frac{\hat{T} - \widehat{T^*}}{2i}(J_o)$ are real, and the latter number is not zero. Thus (2) is established.

We will now prove that $\Delta(B) = C(M)$. The conditions (a), (b), and (d) (Theorem III.3.8) in Theorem IV.2.1 are clearly satisfied by $\Delta(B)$ as a subset of $C(M)$. Consider the linear space $C(M)$ with the supremum norm, $\|\cdot\|_\infty$, defined by

$$\|f\|_\infty = \sup\{|f(J)| : J \in M\}$$

for every $f \in C(M)$. By (1), it follows that

$$\|\hat{T}\|_\infty = \|T\| \text{ for every } T \in B.$$

This means that the isomorphism $\Delta : T \to \hat{T}$ is also an isometry. Thus, for any sequence $\{\hat{T}_n\}$ in $\Delta(B)$ with uniform limit $f_\infty \in C(M)$, we have that $s - \lim_{n \to \infty} T_n = T$ exists in B (by completeness of B). Therefore, $f_\infty =$

$\hat{T} \in \Delta(B)$. By part (2) of this theorem, the functions $(\hat{T} + \widehat{T^*})$ and $-i(\hat{T} - \widehat{T^*})$ are real-valued on M. Hence $\widehat{T^*}(J) = \overline{\hat{T}(J)}$ on M, as is seen by using the Hermitian decomposition of T, and taking the conjugate of each side. Since B is closed under formation of adjoints, $\Delta(B)$ contains $\overline{\hat{T}}$ whenever it contains $\hat{T}$. Therefore, all of the conditions (a), (b), (c), (d), and (e) in Theorem IV.2.1 are satisfied. Consequently the image $\Delta(B)$ of the Gelfand representation is equal to the space of <u>all</u> continuous, complex-valued functions on M.

It remains to prove (3). Let $\hat{T}(J) \geq 0$ on M. Then the function $\hat{S}$ on M defined by

$$\hat{S}(J) = (\hat{T}(J))^{\frac{1}{2}}, \quad J \in M$$

is continuous. By the isomorphism $\Delta : T \to \hat{T}$, there must exist an operator $S \in B$ such that $S^2 = T$. Moreover, $S = S^*$ since $\hat{S}(J) \geq 0$ on M, by (2). Thus

$$(Tx, x) = (S^2x, x) = (Sx, Sx) \geq 0,$$

for every $x \in H$. As stated in Remark II.3.1, this implies that T is Hermitian and positive. Put $\hat{T}_1(J) = \max\{\hat{T}(J), 0\}$ and $\hat{T}_2(J) = \hat{T}_1(J) - \hat{T}(J)$, for every $J \in M$. Then $\hat{T}_1$ and $\hat{T}_2$ are nonnegative, continuous functions on M. It follows by the above remarks that there exist positive operators $T_1, T_2 \in B$ corresponding to the functions $\hat{T}_1, \hat{T}_2$ with $(T_jx, x) \geq 0$ for every $x \in H$ $(j = 1, 2)$. Since

$$\widehat{T_1T_2}(J) = \hat{T}_1(J)\,\hat{T}_2(J) = \begin{cases} \hat{T}(J)\cdot 0 & \text{if } \hat{T}(J) \geq 0 \\ 0\cdot(-\hat{T}(J)) & \text{if } \hat{T}(J) < 0 \end{cases},$$

it follows that $T_1T_2 = \mathbf{0}$. We also have that $T_2 = T_1 - T$ and hence $T_2T_2 + T_2T = \mathbf{0}$. Therefore, for every $x \in H$,

$$0 \leq (T(T_2x), T_2x) = ((T_1 - T_2)(T_2x), T_2x)$$

$$= (-T_2T_2x, T_2x) = -(T_2^3x, x),$$

since T_2 is Hermitian. Thus $(T_2^3x, x) \leq 0$ for every $x \in H$. On the other hand, again since T_2 is Hermitian,

$$(T_2^3x, x) = (T_2(T_2x), T_2x) \geq 0 \quad .$$

Thus $(T_2^3x, x) = 0$ for every $x \in H$. It follows, by Theorem II.3.2, that $\|T_2^3\| = 0$ and thus $T_2^3 = \mathbf{0}$, the zero operator on H. Since $\|T_2\| = \lim_{n \to \infty} \|T_2^n\|^{\frac{1}{n}}$, by Theorem III.3.11, it follows that $T_2 = \mathbf{0}$. But then $\mathbf{0} = T_1 - T$, and hence $T = T_1$. Consequently T is Hermitian, and thus $\hat{T}(J) \geq 0$ for every $J \in M$.

3. Strong Limits of Monotone Sequences of Positive Operators

Theorem 1. Let $\{T_n\}$ be a sequence of Hermitian operators in $B = N(T_o)''$ such that

$$0 \le T_1 \le T_2 \le \ldots \le T_n \le \ldots \le S \in B . \qquad (1)$$

Then $s - \lim_{n \to \infty} T_n x = T_x$ exists for every $x \in H$, that is, $s - \lim_{n \to \infty} T_n = T$ exists, and $T \in B$ with $T_n \le T \le S$ $(n = 1, 2, 3, \ldots)$.

Proof. It is convenient to make use of the generalized Schwarz inequality which asserts that, if T is a positive Hermitian operator on H, then

$$|(Tx, y)|^2 \le (Tx, x)\,(Ty, y) .$$

This inequality is proved in the following way. If

$$z = x + \lambda(Tx, y)y ,$$

then $z \in H$ and $(Tz, z) \ge 0$ for all real scalars λ, since $T \ge 0$. A simple calculation (which uses the fact that T is Hermitian) shows that

$$(Tz, z) = (Tx, x) + 2\lambda|(Tx, y)|^2 + \lambda^2 |(Tx, y)|^2 (Ty, y) .$$

Since $(Tz, z) \ge 0$ for all real λ, it follows that the discriminant is non-positive, and the generalized Schwarz inequality follows from this observation. The operator $R_{nk} = T_{n+k} - T_n$ (where $n \ge 1$, $k \ge 1$) is positive and Hermitian, and $R_{nk} \in B$. Since

$$(T_n x, x) \le (Sx, x) \le \|S\| \text{ if } \|x\| \le 1 ,$$

it follows by Theorem II.3.2 that $\|T_n\| \le \|S\|$ for $n = 1, 2, \ldots$. Consequently, since $T_n \le T_{n+k}$, $\|T_{n+k} - T_n\| \le \|S\|$, also. Thus $\|R_{nk}\| \le \|S\|$ for all $n \ge 1$ and $k \ge 1$. By the generalized Schwarz inequality, for each $x \in H$ we have

$$\|R_{nk}x\|^4 = (R_{nk}x, R_{nk}x)^2 \le (R_{nk}x, x)\ (R_{nk}^2 x, R_{nk}x) \quad .$$

By the usual Schwarz inequality, and the fact that

$$\|R_{nk}x\| \le \|R_{nk}\| \cdot \|x\| \le \|S\| \cdot \|x\| \quad ,$$

we find that

$$\|R_{nk}x\|^4 \le (R_{nk}x, x)\ \|S\|^3\ \|x\|^2 \quad .$$

Hence, since $R_{nk} = T_{n+k} - T_n$,

$$\|T_{n+k}x - T_n x\|^4 \le ((T_{n+k}x, x) - (T_n x, x))\ \|S\|^3 \|x\|^2 \quad .$$

Since the sequence $\{(T_n x, x)\}$ is a convergent increasing sequence of positive numbers, it follows that $\{T_n x\}$ is a Cauchy sequence in the norm of the Hilbert space H. Since H is complete, there exists a limit Tx, for each

$x \in H$. Thus $s - \lim_{n \to \infty} T_n = T$. Since the algebra B is a Banach algebra, it follows that $T \in B$. Note that, by continuity of the inner product in its first argument,

$$(Tx, x) = \lim_{n \to \infty} (T_n x, x) \geq 0$$

for all $x \in H$. Thus T is Hermitian. In addition, it follows easily that $T_n \leq T \leq S$ for $n = 1, 2, \ldots$.

Definition 1. A set E in a topological space X is said to be nowhere dense in X if the closure $\bar{E}$ contains no nonempty open sets of X. We say that a set E in X is of the first category in X if E is expressible as the union of a countable number of sets each of which is nowhere dense in X; otherwise E is of the second category in X.

Remark 1. We shall mention here two important theorems due to Baire. Proofs of these theorems are given, for example, in Yosida [30], p. 12.

Baire's Theorem 1. Let D be a set of the first category in a compact topological space X. Then the complement $D^c = X - D$ is dense in X.

Baire's Theorem 2. Let $\{x_n(t)\}$ be a sequence of real-valued continuous functions defined on a topological space X. Suppose that a finite limit:

$$\lim_{n \to \infty} x_n(t) = x(t)$$

exists at every point t of X. Then the set of points at which the function x is discontinuous is a set of the first category in X.

<u>Theorem 2</u>. Let a sequence of real-valued functions $\{\hat{T}_n\}$ in C(M) satisfy the condition

$$0 \leq \hat{T}_1(J) \leq \hat{T}_2(J) \leq \ldots \leq \hat{T}_n(J) \leq \ldots \leq \alpha \qquad (2)$$

for every $J \in M$, where M is the structure space of $B = N(T_o)''$ (α being a finite constant). Then $s - \lim_{n \to \infty} T_n = T$ exists, the set

$$D = \{J : J \in M, \hat{T}(J) \neq \lim_{n \to \infty} \hat{T}_n(J)\}$$

is of the first category in M, and $D^c = M - D$ is dense in the space M.

<u>Proof</u>. From the hypothesis (2) and part (3) of Theorem IV.2.2 we have

$$0 \leq T_1 \leq T_2 \leq \ldots \leq T_n \leq \ldots \leq \alpha I \in B,$$

and $\{T_n\}$ is a sequence of Hermitian positive operators in B. By Theorem IV.3.1, $s - \lim_{n \to \infty} T_n = T$ exists, $T \in B$, and $T_n \leq T \leq \alpha I$ $(n \geq 1)$. Put $F(J) = \lim_{n \to \infty} \hat{T}_n(J)$ on M; F(J) is the limit of the nondecreasing sequence $\{\hat{T}_n(J)\}$ bounded above by α, for each $J \in M$. By the above remarks, $F(J) \leq \hat{T}(J)$ on M. Suppose, to the contrary of the conclusion, that the set $D = \{J : J \in M, \hat{T}(J) \neq F(J)\}$ is of the second category in M. Since the discontinuity set D_F of F is of the first category by Baire's Theorem 2, and the set D is assumed to be of second category, there exists a point $J_o \in D$ such that F is continuous at J_o. Otherwise we would have $D \subseteq D_F$, which would imply that D_F is of second category. Since $F(J) \le \hat{T}(J)$ on M, and $F(J_o) \ne \hat{T}(J_o)$, it follows that $F(J_o) < \hat{T}(J_o)$. The nonnegative function defined on M by $\hat{T}(J) - F(J)$ is strictly positive at J_o. Thus there exists a number $\delta > 0$ and an open set $V(J_o)$ containing J_o such that

$$\hat{T}(J) - F(J) \ge \delta \text{ for every } J \in V(J_o) \quad . \tag{3}$$

Now M is a compact Hausdorff space, and there exists an open set $V_1 \subset M$ with compact closure such that

$$\{J_o\} \subset V_1 \subset \overline{V}_1 \subset V(J_o) \quad .$$

This fact is proved in Rudin [25], p. 37, for example. By Urysohn's lemma (see, for example, Kelley [13], p. 115) and the fact that $\overline{V}_1$ and $V(J_o)^c$ are disjoint closed sets, there exists a function $\hat{G} \in C(M)$ such that $0 \le \hat{G}(J) \le \delta$ on M, $\hat{G}(J) = \frac{1}{2}\delta$ on $\overline{V}_1$ (and thus certainly on V_1), and $\hat{G}(J) = 0$ on $V(J_o)^c$. It follows from (3) that

$$F(J) + \hat{G}(J) \le \hat{T}(J) \text{ for all } J \in M \quad .$$

This implies that $T_n \le T - G$ ($n \ge 1$), where G is the operator in B corresponding to the representative function $\hat{G} \in C(M)$. Since $\hat{G}(J) \ge 0$, it follows by part (3) of Theorem IV.2.2 that the operator $G \ge 0$. However $G \ne 0$, since $\hat{G}(J) > 0$ on V_1. Thus we have $s - \lim_{n \to \infty} T_n \le T - G$ with $G \ge 0$, $G \ne 0$, and this contradicts the fact that $T = s - \lim_{n \to \infty} T_n$. Thus D must be a set of the first category. Since M is compact, the complement of the set D of the first category in M is dense in M by Baire's Theorem 1.

4. The Spectral Resolution of a Normal Operator

In this concluding section we shall show how the Gelfand theory of Banach algebras can be applied to obtain a spectral resolution (or spectral decomposition) of a normal operator on the Hilbert space H.

In what follows, given an arbitrary but fixed normal operator $T_o \in A(H)$, we shall denote by $B = N(T_o)''$, the commutative Banach algebra with unit I generated by the normal operator T_o, and M will denote the structure space of B.

Remark 1. Let $C'(M)$ denote the set of all complex-valued bounded functions f on M which differ from a continuous function $\hat{T}$ only on a set of the first category in M. Define the relation $\simeq$ on $C'(M)$ as follows. If f, g, $\in C'(M)$, then $f \simeq g$ if and only if the set

$$\mathcal{E}_{f,g} = \{J : J \in M,\ f(J) \neq g(J)\} \tag{1}$$

is a set of the first category in M. The equivalence realtion $\simeq$ partitions $C'(M)$ into distinct equivalence classes which will be denoted by

$$(f) = \{g : g \in C'(M),\ g \simeq f\} \tag{2}$$

for every $f \in C'(M)$.

Remark 2. Suppose that $\hat{T}_1, \hat{T}_2 \in C(M)$ and $\hat{T}_1 \neq \hat{T}_2$, that is, for some $J_o \in M$, $\hat{T}_1(J_o) \neq \hat{T}_2(J_o)$. By continuity of the functions $\hat{T}_1$, $\hat{T}_2$ on M, there exists an open set $U(J_o) \subset M$ containing J_o such that, for each

$J \in U(J_o)$, $\hat{T}_1(J) \neq \hat{T}_2(J)$. Clearly, $U(J_o) \subset \mathcal{E}_{\hat{T}_1, \hat{T}_2}$. Since the complement of a set of the first category in the compact Hausdorff space M is dense in M, $\mathcal{E}_{\hat{T}_1, \hat{T}_2}$ cannot be of the first category in M. Thus $\hat{T}_1 \notin (\hat{T}_2)$. The above remarks prove that each of the classes (f) contains exactly one continuous function $\hat{T}$. Therefore the set of all the equivalence classes (f) is in one-to-one correspondence with the elements of the Banach algebra B by the Gelfand representation $\Delta : T \to \hat{T}$.

Remark 3. Consider a fixed operator $T \in B$, and let λ be a complex number. Define

$$Q_\lambda = \{J : J \in M, \operatorname{Re}\hat{T}(J) < \operatorname{Re}\lambda, \operatorname{Im}\hat{T}(J) < \operatorname{Im}\lambda\} \quad , \tag{3}$$

and let $E_\lambda' = \chi_{Q_\lambda}$, the characteristic function of the set Q_λ. It will be shown that $E_\lambda' \in C'(M)$. For any complex number ξ, let

$$\Gamma_\xi = \{z : z \in C, \operatorname{Re}z < \operatorname{Re}\xi, \operatorname{Im}z < \operatorname{Im}\xi\} \quad . \tag{4}$$

There exists an increasing sequence $\{f_n\}$ of continuous real-valued functions defined on the complex plane such that

$$\lim_{n \to \infty} f_n(z) = \chi_{\Gamma_\lambda}(z)$$

for every complex number z. Such a sequence can be defined directly in the present case, but existence of such a sequence follows from the fact that χ_{Γ_λ} is lower semicontinuous on the complex plane (as the characteristic function of an open set) and by a standard theorem (see, for example, McShane and Botts [17], p. 76) there exists a sequence $\{f_n\}$ of continuous real-valued functions such that $f_n(z) \uparrow \chi_{\Gamma_\lambda}(z)$ for every complex number z. It is proved in the reference cited that an appropriate sequence is defined by

$$f_n(z) = \inf\{\chi_{\Gamma_\lambda}(w) + n|w - z| : w \in C\} , \tag{5}$$

(n = 1, 2, ...). Thus

$$E_\lambda'(J) = \chi_{\Gamma_\lambda}(\hat{T}(J)) = \lim_{n \to \infty} f_n(\hat{T}(J)) \tag{6}$$

for every $J \in M$. By Baire's Theorem 2, as referred to in Section 3 of this chapter, the set of discontinuities of E_λ' is of the first category in the space M. Consequently $E_\lambda' \in C'(M)$, by definition of the set $C'(M)$.

Remark 4. Consider the equivalence class (E_λ') containing the characteristic function E_λ'. Let E_λ be the operator in B which corresponds to the equivalence class (E_λ') by the Gelfand representation $\Delta : T \to \hat{T}$ (as in Remark 2). The operators $E_\lambda \in B$ so obtained will be used to construct a type of spectral resolution of the normal operator $T \in B$, as follows.

Let $\varepsilon > 0$ be given. For the fixed normal operator $T \in B$, let

$$\alpha = \sup\{|\mathrm{Re}\ \hat{T}(J)| : J \in M\}, \quad \text{and} \tag{7}$$

$$\beta = \sup\{|\mathrm{Im}\ T(J)| : J \in M\} . \tag{8}$$

Note that the complex numbers $-\alpha - i\beta$, $-\alpha + i\beta$, $\alpha + i\beta$, $\alpha - i\beta$ are the vertices of a rectangle R which circumscribes the spectrum of T. Let

$$-\alpha = \alpha_1 < \alpha_2 < \ldots < \alpha_n = \alpha + \frac{\varepsilon}{\sqrt{2}} \tag{9}$$

and

$$-\beta = \beta_1 < \beta_2 < \ldots < \beta_n = \beta + \frac{\varepsilon}{\sqrt{2}} \tag{10}$$

be partitions of the intervals $[-\alpha, \alpha + \frac{\varepsilon}{\sqrt{2}}]$, $[-\beta, \beta + \frac{\varepsilon}{\sqrt{2}}]$ respectively, such that

$$\{\max_{1\le j\le n} (\alpha_j - \alpha_{j-1})^2 + \max_{1\le j\le n} (\beta_j - \beta_{j-1})^2\}^{\frac{1}{2}} \le \varepsilon \quad . \tag{11}$$

The half-open cells of the form

$$R_{jk} = \{\zeta + i\eta : \alpha_{j-1} \le \zeta < \alpha_j, \beta_{k-1} \le \eta < \beta_k\} \tag{12}$$

for $j = 2, \ldots, n$, $k = 2, \ldots, n$ are pairwise disjoint and cover the closed rectangle R which contains the spectrum of T. Let ζ_j, η_k for $j = 1, 2, \ldots, n$, $k = 1, 2, \ldots, n$ be real numbers such that $\zeta_j + i\eta_k \in R_{j+1\ k+1}$ for $j = 1, 2, \ldots, n$, $k = 1, 2, \ldots, n$. We shall first prove that for every $J \in M$,

$$|\hat{T}(J) - \sum_{j=2}^{n} \sum_{k=2}^{n} (\zeta_j + i\eta_k) E''_{\zeta_j,\eta_k}(J)| \le \epsilon, \tag{13}$$

where

$$E''_{\zeta_j,\eta_k}(J) = E'_{\zeta_j+i\eta_k}(J) + E'_{\zeta_{j-1}+i\eta_{k-1}}(J) - E'_{\zeta_j+i\eta_{k-1}}(J)$$

$$- E_{\zeta_{j-1}+i\eta_k}(J) \quad .$$

Recall that the range of $\hat{T}$ coincides with the spectrum of T (Theorem III. 3.10). Hence if $J \in M$, then there exist integers r,s ($2 \le r \le n$, $2 \le s \le n$) such that $\hat{T}(J) \in R_{rs}$. For this J, we have

$$\sum_{j=2}^{n} \sum_{k=2}^{n} (\zeta_j + i\eta_k) E''_{\zeta_j,\eta_k}(J) = \zeta_r + i\eta_s$$

since the half-open cells $\{R_{jk} : j = 2, \ldots, n, k = 2, \ldots, n\}$ are pair-wise disjoint. Therefore, by (11)

$$|\hat{T}(J) - (\zeta_r + i\eta_s)| \le \epsilon \quad .$$

This proves (13). Now, let $\hat{E}_\lambda$ be the continuous function defined on M corresponding to the operator $E_\lambda \in B$ (or, corresponding to the equivalence class (E'_λ), as in Remark 2). By definition, $\hat{E}_\lambda$ differs from the characteristic function E'_λ only on a set $\mathcal{E}_{\hat{E}_\lambda, E'_\lambda}$ of the first category in M. By

Baire's Theorem 1, the complement, $\mathcal{E}^{c}_{\hat{E}_\lambda, E_\lambda'}$, of $\mathcal{E}_{\hat{E}_\lambda, E_\lambda'}$ is dense in the structure space M. It follows from (13) that

$$|\hat{T}(J) - \sum_{j=2}^{n} \sum_{k=2}^{n} \lambda_{j,k} \hat{\hat{E}}_{\lambda_{j,k}}(J)| \leq \epsilon \quad , \tag{14}$$

where $\lambda_{j,k} = \zeta_j + i\eta_k$ for j = 2, ..., n, k = 2, ..., n, and

$$\hat{\hat{E}}_{\lambda_{j,k}}(J) = \hat{E}_{\lambda_{j,k}}(J) + \hat{E}_{\lambda_{j-1,k-1}}(J) - \hat{E}_{\lambda_{j,k-1}}(J) - \hat{E}_{\lambda_{j-1,k}}(J) \quad .$$

By part (1) of Theorem IV.2.2, we have

$$\|T - \sum_{j=2}^{n} \sum_{k=2}^{n} \lambda_{j,k} \left[E_{\lambda_{j,k}} + E_{\lambda_{j-1,k-1}} - E_{\lambda_{j-1,k}} - E_{\lambda_{j,k-1}}\right]\| \leq \epsilon \quad . \tag{15}$$

Since $\epsilon > 0$ was arbitrary in the above argument, there corresponds to every $\epsilon > 0$ a partition $\{R_{jk}: j = 2, \ldots, n, k = 2, \ldots, n\}$ of the closed rectangle R containing $\sigma(T)$ such that (15) is true for every choice of the complex numbers $\lambda_{j,k} \in R_{jk}$ ($j = 2, \ldots, n, k = 2, \ldots, n$). We shall denote this result by

$$T = \iint \lambda \, dE_\lambda \quad , \tag{16}$$

and (16) is called the <u>spectral</u> <u>resolution</u> of the normal operator T in the sense of Yosida.

<u>Note</u>. The fact that every bounded normal operator on a Hilbert

space has a spectral resolution is known as _the spectral theorem_. The proof of the spectral theorem given above is a much expanded version of the proof due to K. Yosida (refer to Yosida [29], or Yosida [30]).

BIBLIOGRAPHY

1. Bachman, G. and Narici, L., Functional Analysis. Academic Press, New York (1966).

2. Berberian, S. K., Introduction to Hilbert Space. Oxford University Press, New York (1961).

3. Bernau, S. J., The spectral theorem for normal operators. Journal of the London Mathematical Society 40 (1965), pp. 478-486.

4. Bernau, S. J. and Smithies, F., A note on normal operators. Proceedings of the Philosophical Society 59 (1963), pp. 727-729.

5. Dunford, N. and Schwartz, J. T., Linear Operators Part II: Spectral Theory. Interscience, New York (1963).

6. Freudenthal, H., Teilweise geordnete Moduln. Akademie Van Wetenschappen, Amsterdam, Proceedings of the Section of Sciences 39 (1936), pp. 641-651.

7. Gelfand, I. M., Normierte Ringe. Matematičheviskiĭ Subornik 9 (1941), pp. 3-24.

8. Gelfand, I. M., Raikov, D. A. and Shilov, G. E., Commutative Normed Rings, Chelsea. New York (1964).

9. Halmos, P. R., Introduction to Hilbert Space and the Theory of Spectral Multiplicity, Chelsea. New York (1951).

10. Hewitt, E. and Stromberg, K., Real and Abstract Analysis, Springer. Berlin (1965).

11. Hille, E. and Phillips, R. S., Functional Analysis and Semi-Groups. American Mathematical Society, Providence. Colloquium Publications, Vol. XXXI (1957).

12. Kakutani, S., Mean ergodic theorem in abstract (L) - space. Imperial Academy of Japan, Tokyo. Proceedings 15 (1939), pp. 121-123.

13. Kelley, J. T., General Topology. D. Van Nostrand, Princeton (1955).

14. Loomis, L. H., An Introduction to Abstract Harmonic Analysis. D. Van Nostrand, Princeton (1953).

15. Lorch, E. R., Spectral Theory. Oxford University Press, New York (1962).

16. Markushevich, A. I., Theory of Functions of a Complex Variable, Vol. I. (trans. R. A. Silverman), Prentice-Hall, Englewood Cliffs (1965).

17. McShane, E. J. and Botts, T. A., Real Analysis. D. Van Nostrand, Princeton (1959).

18. Nagumo, M., Einige analytische Untersuchungen in linearen metrischen Ringen. Japanese Journal of Mathematics 13 (1936), pp. 61-80.

19. Naimark, M. A., Normed Rings (trans. L. F. Boron), P. Noordhoff. Groningen (1959).

20. Neumann, J. von, On an abstract generalization of the quantum mechanical formulation (part I). Matematičheskiĭ Sbornik 43 (1936), pp. 415-482.

21. Neumann, J. von, Über adjungierte Funktionaloperatoren. Annals of Mathematics 33 (1932), pp. 294-310.

22. Rickart, C. E., General Theory of Banach Algebras. D. Van Nostrand, Princeton (1960).

23. Riesz, F., Sur quelques notions fondamentales dans la théorie générale des operations lineaires. Annals of Mathematics 41 (1940), pp. 174-206.

24. Royden, H. L., Real Analysis. Macmillan, New York (1963).

25. Rudin, W., Real and Complex Analysis. McGraw-Hill, New York (1966).

26. Steen, S., An introduction to the theory of operators (I). Proceedings of the London Mathematical Society 40 (1936), pp. 361-392.

27. Stone, M. H., A general theory of spectra (I). National Academy of Sciences, Washington. Proceedings 26 (1940), pp. 280-283; (II), ibid: 37 (1941), pp. 83-87.

28. Taylor, A. E., Introduction to Functional Analysis. Wiley, New York, (1958).

29. Yosida, K., Normed rings and spectral theorems. Proceedings of the Imperial Academy of Japan, Tokyo 19 (1943), pp. 356-359.

30. Yosida, K., Functional Analysis. Springer, Berlin (1965).

31. Yosida, K., On the theory of spectra. Proceedings of the Imperial Academy of Japan, Tokyo. Proceedings 16 (1940), pp. 378-383.

32. Yosida, K., and Nakayama, T., On the semi-ordered ring and its application to the spectral theorem (I). Proceedings of the Imperial Academy of Japan, Tokyo. Proceedings 18 (1942), pp. 144-147.

33. Vulich, B., Une définition du produit dans les espaces semi-ordonnés linéaires. Dopovidi Akademiï Nauk Ukraïns'koï RSR (Kiev) 26 (1940), pp. 850-859.