MANONMANIAM SUNDARANAR UNIVERSITY

TIRUNELVELI-627 012

DIRECTORATE OF DISTANCE AND CONTINUING EDUCATION

II M.Sc. MATHEMATICS

SEMESTER IV

CORE XI : FUNCTIONAL ANALYSIS

Sub. Code: SMAM41

Prepared by

Dr. Leena Nelson S N

Associate Professor & Head, Department of Mathematics

Women's Christian College, Nagercoil – 1.

FUNCTIONAL ANALYSIS (SMAM41)

UNIT	DETAILS
I	Banach Spaces: The definition and some examples – Continuous linear transformations – The Hahn-Banach theorem – The natural imbedding of N in N**.
II	The open mapping theorem – The conjugate of an Operator. The definition and some simple properties – Orthogonal complements – Orthonormal sets.
III	The conjugate space H* - The adjoint of an operator – self-adjoint operators – Normal and Unitary operators – Projections.
IV	Finite-Dimensional Spectral Theory: Determinants and the spectrum of an operator – The spectral theorem.
V	General Preliminaries on Banach Algebras: The definition and some examples – Regular and singular elements – Topological divisors of zero – The spectrum – The formula for the spectral radius – The radical and semi-simplicity.

Text Book
G.F. Simmons, Introduction to Topology and Modern Analysis, McGraw Hill Education (India) Private Limited, New Delhi, 1963.

UNIT – 1

1.1. THE DEFINITION AND SOME EXAMPLES

We begin by restating the definition of a Banach space.

A normed linear space is a linear space N in which to each vector x there corresponds a real number, denoted by $\|x\|$ and called the norm of x, in such a manner that

(1) $\|x\| > 0$, and $\|x\| = 0 \Leftrightarrow x = 0$;
(2) $\|x + y\| \leq \|x\| + \|y\|$;
(3) $\|\alpha x\| = |\alpha|\|x\|$,

The non-negative real number $\|x\|$ is to be thought of as the length of the vector x. If we regard $\|x\|$ as a real function defined on N, this function is called the norm on N. It is easy to verify that the normed linear space N is a metric space with respect to the metric d defined by $d(x, y) = \|x - y\|$. A Banach space is a complete normed linear space. Our main interest in this chapter is in Banach spaces, but there are several points in the body of the theory at which it is convenient to have the basic definitions and some of the simpler facts formulated in terms of normed linear spaces. For this reason, and also to emphasize the role of completeness in theorems which require this assumption, we work in the more general context whenever possible. The reader will find that the deeper theorems, in which completeness hypotheses are necessary, often make essential use of Baire's theorem.

Several simple but important facts about a normed linear space are based on the following inequality:

$$|\|x\| - \|y\|| \leq \|x - y\| \tag{1}$$

To prove this, it suffices to prove that

$$\|x\| - \|y\| \leq \|x - y\| \tag{2}$$

for it follows from (2) that we also have

$$-(\|x\| - \|y\|) = \|y\| - \|x\| \leq \|y - x\| = \|-(x - y)\| = \|x - y\|,$$

which together with (2) yields (1). We now prove (2) by observing that $\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|$. The main conclusion we draw from (1) is that the norm is a continuous function:

$$x_n \to x \Rightarrow \|x_n\| \to \|x\|.$$

This is clear from the fact that $|\|x_n\| - \|x\|| \leq \|x_n - x\|$, since $x_n \to x$ means that $\|x_n - x\| \to 0$. In the same vein, we can prove that addition and scalar multiplication are jointly continuous (see Problem 22-5), for

$$x_n \to x \text{ and } y_n \to y \Rightarrow x_n + y_n \to x + y$$

$$\alpha_n \to \alpha \text{ and } x_n \to x \Rightarrow \alpha_n x_n \to \alpha x$$

These assertions follow from

$$\|(x_n + y_n) - (x + y)\| = \|(x_n - x) + (y_n - y)\| \leq \|x_n - x\| + \|y_n - y\|$$

and

$$\|\alpha_n x_n - \alpha x\| = \|\alpha_n (x_n - x) + (\alpha_n - \alpha)x\| \leq |\alpha_n| \|x_n - x\| + |\alpha_n - \alpha| \|x\|.$$

Our first theorem exhibits one of the most useful ways of forming new normed linear spaces out of old ones.

Theorem 1.1 : Let M be a closed linear subspace of a normed linear space N. If the norm of a coset $x + M$ in the quotient space $\frac{N}{M}$ is defined by

$$\|x + M\| = \inf\{\|x + m\| : m \in M\}, \qquad (3)$$

then $\frac{N}{M}$ is a normed linear space. Further, if N is a Banach space, then so is $\frac{N}{M}$.

Proof.

We first verify that (3) defines a norm in the required sense. It is obvious that $\|x+M\| \geq 0$; and since M is closed, it is easy to see that $\|x+M\| = 0 \Leftrightarrow$ there exists a sequence $\{m_k\}$ in M such that $\|x+m_k\| \to 0 \Leftrightarrow x$is in $M \Leftrightarrow x+M = M$ = the zero element of$\frac{N}{M}$. Next, we have

$$\|(x+M)+(y+M)\| = \|(x+y)+M\|$$

$$= \inf\{\|x+y+m\| : m \in M\}$$
$$= \inf\{\|x+y+m+m'\| : m \,\&m' \in M\}$$

$$= \inf\{\|(x+m)+(y+m')\| : m \,\&m' \in M\}$$

$$\leq \inf\{\|x+m\| + \|y+m'\| : m \,\&m' \in M\}$$

$$= \inf\{\|x+m\| : m \in M\} + \inf\{\|y+m'\| : m' \in M\}$$

$$= \|x+M\| + \|y+M\|.$$

The proof of $\|\alpha(x+M)\| = |\alpha|\|x+M\|$ is similar.

Finally, we assume that N is complete, and we show that $\frac{N}{M}$ is also complete. If we start with a Cauchy sequence in $\frac{N}{M}$, then by Problem 12-2 it suffices to show that this sequence has a convergent sub- sequence.

It is clearly possible to find a subsequence $\{x_n + M\}$ of the original Cauchy sequence such that $\|(x_1+M)-(x_2+M)\| < \frac{1}{2}$, $\|(x_2+M)-(x_3+M)\| < \frac{1}{4}$, and, in general, $\|(x_n+M)-(x_{n+1}+M)\| < \frac{1}{2^n}$.

We prove that this sequence is convergent in $\frac{N}{M}$. We begin by choosing any vector y_1 inx_1+M, and we select y_2 inx_2+M such that$\|y_1-y_2\| < \frac{1}{2}$.

We next select a vector y_3 in $x_3 + M$ such that $\|y_2 - y_3\| < \frac{1}{4}$. Continuing in this way, we obtain a sequence $\{y_n\}$ in N such that $\|y_n - y_{n+1}\| < \frac{1}{2^n}$. If $m < n$, then

$$\|y_m - y_n\| = \|(y_m - y_{m+1}) + (y_{m+1} + y_{m+2}) + \cdots + (y_{n-1} - y_n)\|$$

$$\leq \|y_m - y_{m+1}\| + \|y_{m+1} + y_{m+2}\| + \cdots + \|y_{n-1} - y_n\|$$

$$< \frac{1}{2^m} + \frac{1}{2^{m+1}} + \cdots + \frac{1}{2^{n-1}}$$

$$< \frac{1}{2^{m-1}}$$

so $\{y_n\}$ is a Cauchy sequence in N. Since N is complete, there exists a vector y in N such that $y_n \to y$. It now follows from $\|(x_n + M) - (y + M)\| \leq \|y_n - y\|$ that $x_n + M \to y + M$, so $\frac{N}{M}$ is complete.

In the following sections and chapters, we shall often have occasion to consider the quotient space of a normed linear space with respect to a closed linear subspace. In accordance with our theorem, a quotient space of this kind can always be regarded as a normed linear space in its own right.

We now describe some of the main examples of Banach spaces. In each of these, the linear operations are understood to be defined either coordinatewise or pointwise, whichever is appropriate in the circumstances.

Example 1. The spaces R and C—the real numbers and the complex numbers—are the simplest of all normed linear spaces. The norm of a number x is of course defined by $\|x\| = |x|$, and each space is a Banach space.

Example 2. The linear spaces R^n and C^n of all n-tuples $x = (x_1, x_2, \ldots, x_n)$ of real and complex numbers can be made into normed linear spaces in an infinite variety of ways, as we shall see below. If the norm is defined by

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \tag{4}$$

then we get the n-dimensional Euclidean and unitary spaces familiar to us from our earlier work. We denoted these spaces by R^n and C^n in Part 1 of the text book, and we know by the theorems of Sec. 15 that both are Banach spaces.

Each of the following examples consists of n-tuples of scalars, sequences of scalars, or scalar-valued functions defined on some non- empty set, where the scalars are the real numbers or the complex numbers. We do not normally specify which system of scalars is to be used, and it should be emphasized that both possibilities are allowed unless the contrary is clearly stated. Also, we make no distinction in notation between the real case and the complex case. When it turns out to be necessary to distinguish these two cases, we do so verbally, by referring, for instance, to ''the complex space —.'' These conventions are in accord with the standard usage preferred by most mathematicians, and they enable us to avoid a good deal of cumbersome notation and many unnecessary case distinctions.

Example 3. Let p be a real number such that$1 \leq p < \infty$. We denote by l_p^n the space of all n-tuples $x = (x_1, x_2, \dots, x_n)$of scalars, with the norm defined by

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \tag{5}$$

Formula (4) is obviously the special case of (5) which corresponds to $p = 2$, so the real and complex spaces l_2^n are the n-dimensional Euclidean and unitary spaces R^n and C^n. It is easy to see that (5) satisfies condi- tions (1) and (3) required by the definition of a norm. In Problem 4 we outline a proof of the fact that (5) also satisfies condition (2), that is, that $\|x + y\|_p \leq \|x\|_p + \|y\|_p$. The completeness of l_p^n follows from substantially the same reasoning as that used in the proof of Theorem 15-A, so l_p^n a Banach space.

Example 4. We again consider a real number p with the property that$1 \leq p < \infty$, and we denote by l_p, the space of all sequences

$$x = (x_1, x_2, \dots, x_n, \dots)$$

of scalars such that $\sum_{n=1}^{\infty} |x_n|^p < \infty$, with the norm defined by

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} \tag{6}$$

The reader will observe that the real and complex spaces l2 are precisely the infinite-dimensional Euclidean and unitary spaces R^∞ and C^∞ defined in Problem 15-4. The proof of the fact that 1, actually is a Banach space requires arguments similar to those used in Problems 15-3 and 15-4.

The Banach spaces discussed in these examples are all special cases of the important L_p, spaces studied in the theory of measure and integration. A detailed treatment of these spaces is outside the scope of this book, but we can describe them loosely as follows. An L_p, space essentially consists of all measurable functions f defined on a measure space X with measure m which are such that $|f(x)|^p$ is integrable, with

$$\|f\|_p = \left(\int |f(x)|^p dm(x) \right)^{\frac{1}{p}} \tag{7}$$

taken as the norm. In order to include the spaces l_n^p and l_p, within the theory of L_p, spaces, we have only to consider the sets $\{1, 2, \dots, n\}$ and $\{1, 2, \dots, n, \dots\}$ as measure spaces in which each point has measure 1, and to regard n-tuples and sequences of scalars as functions defined on these sets. Since integration is a generalized type of summation, formulas (5) and (6) are special cases of formula (7).

Example 5. Just as in Example 3, we start with the linear space of all n-tuples $x = (x_1, x_2, \dots, x_n)$of scalars, but this time we define the norm by

$$\|x\| = \max\{|x_1|, |x_2|, \dots, |x_n|\} \tag{8}$$

This Banach space is commonly denoted byl_m^n, and the symbol$\|x\|_\infty$. is occasionally used for the norm given by (8). The reason for this practice lies in the interesting fact that

$$\|x\|_\infty = \lim\|x\|_p \quad as \;\; p \to \infty$$

that is, that

$$\max\{|x_i|\} = \lim\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}} \quad as \quad p \to \infty \tag{9}$$

We briefly inspect the case $n = 2$ to see why thisistrue. Let$x = (x_1, x_2)$ be an ordered pair of real numbers with x_1 and$x_2 \geq 0$. It is clear that$\|x\|_\infty = \max\{x_1, x_2\} \leq \left(x_1^p + x_2^p\right)^{\frac{1}{p}} = \|x\|_p$. If$x_1 = x_2$, thenlim$\|x\|_p = \lim\left(2x_2^p\right)^{\frac{1}{p}} = \lim 2^{\frac{1}{p}}x_2 = \|x\|_\infty$. And if$x_1 < x_2$, then

$$\lim\|x\|_p = \lim\left(x_1^p + x_2^p\right)^{\frac{1}{p}}$$

$$= \lim\left(\left[\left(\frac{x_1}{x_2}\right)^p + 1\right]x_2^p\right)^{\frac{1}{p}}$$

$$= \lim\left[\left(\frac{x_1}{x_2}\right)^p + 1\right]^{\frac{1}{p}} x_2$$

$$= x_2$$

$$= \|x\|_\infty$$

Example 6. Consider the linear space of all bounded sequences $x = \{x_1, x_2, .., x_n, ...\}$of scalars. By analogy with Example 5, we define the norm by

$$\|x\| = \sup|x_n| \tag{10}$$

and we denote the resulting Banach space by l_∞. The set c of all convergent sequences is easily seen to be a closed linear subspace of l_∞. and is therefore itself a Banach space. Another Banach space in this family is the subset c_o of c which consists of all convergent sequences with limit 0.

Example 7. The Banach space of primary interest to us is the space $\mathcal{C}(X)$ of all bounded continuous scalar-valued functions defined on a topological space X, with the norm given by

$$\|f\| = \sup|f(x)| \tag{11}$$

This norm is sometimes called the uniform norm, because the statement that f_n, converges to f with respect to this norm means that f_n, converges to f uniformly on X. The fact that this space is complete amounts to the fact that if f is the uniform limit of a sequence of bounded continuous functions, then f itself is bounded and continuous. If, as above, we consider n-tuples and sequences as functions defined on $\{1, 2, \dots, n\}$ and $\{1,2, \dots, n, \dots\}$, then the spaces $l_\infty^n$and l_∞, are the special cases of $\mathcal{C}(X)$ which correspond to choosing X to be the sets just mentioned, each with the discrete topology.

Many important properties of a Banach space are closely linked to the shape of its closed unit sphere, that is, the set$S = \{x: \|x\| \leq 1\}$. One basic property of S is that it is always conver, in the sense (see Problem 32-5) that if x and y are any two vectors in S, then the vector $z = \alpha x + \beta y$ is also in S, where α and β are non-negative real numbers such that$\alpha + \beta = 1$; for

$$\|z\| = \|\alpha x + \beta y\|$$

$$\leq \alpha\|x\| + \beta\|y\|$$

$$< \alpha + \beta$$

$$= 1$$

In this connection, it is illuminating to consider the shape of S for certainsimple examples. Let our underlying linear space be the real linear

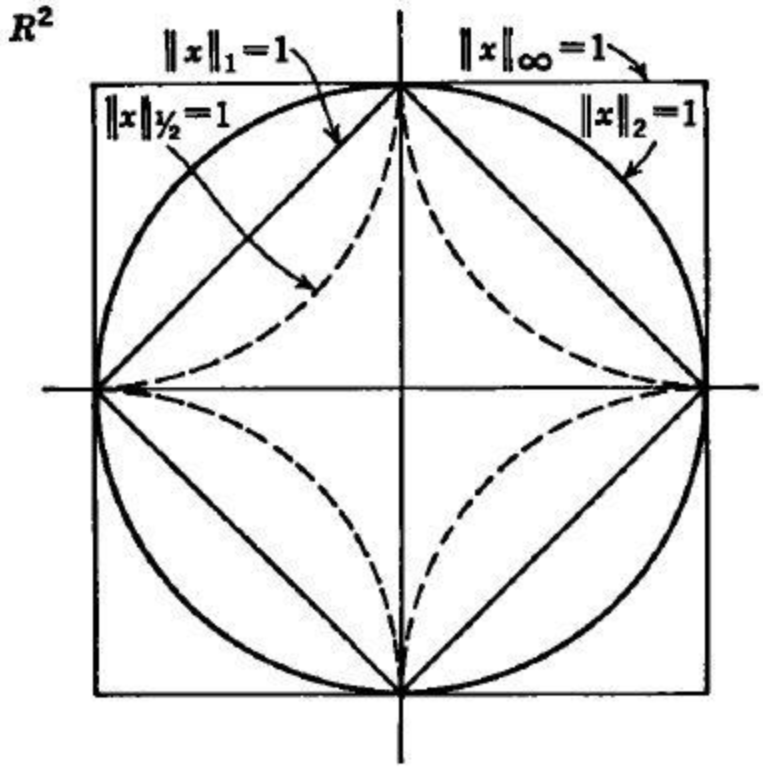

Fig. 1 Some Closed Unit Spheres

space R^2 of all ordered pairs $x = (x_1, x_2)$ of real numbers. As we have seen, there are many different norms which can be defined on R^2, among which are the following:

$$\|x\|_1 = |x_1| + |x_2|; \|x\|_2 = (|x_1|^2 + |x_2|^2)^{\frac{1}{2}} \text{and } \|x\|_\infty = \max\{x_1, x_2\}.$$

Figure 1 illustrates the closed unit sphere which corresponds to each of these norms. In the first case, S is the square with vertices $(1,0), (0,1), (-1,0), (0,-1)$; in the second, it is the circular disc of radius 1; and in the third, it is the square with vertices $(1,1), (-1,1), (-1,-1), (1,-1)$. If we consider the norm defined by

$$\|x\|_p = (|x_1|^p + |x_2|^p)^{\frac{1}{p}} \qquad (12)$$

Where$1 \leq p < \infty$, and if we allow p to increase from 1 to ∞, then the corresponding S's swell continuously from the first square mentioned to the second. We note that S is truly "spherical"$\Leftrightarrow p = 2$. These considerations also show quite clearly why we always assume that $p \geq 1$; for if we were to define $\|x\|_p$, by formula (12) with $p < 1$, then $S = \{x: \|x\|_p \leq 1\}$would not be convex (see the star-shaped inner portion of Fig. 35). For $p < 1$, therefore, formula (12) does not yield a norm.

In the above examples, we have exhibited several different types of Banach spaces, and there are yet others which we have not mentioned. Amid this diversity of possibilities, it is well to realize that any Banach space can be regarded—from the point of view of its linear and norm structures alone—as a closed linear subspace of $\mathcal{C}(X)$ for a suitable compact Hausdorff space X. We prove this below, in our discussion of the natural imbedding of a Banach space in its second conjugate space.

Problems

1. Let N be a non-zero normed linear space, and prove that N is a Banach space $\Leftrightarrow$ $\{x: \|x\| = 1\}$ is complete.
2. Leta Banach space B be the direct sum of the linear subspaces M and N,sothat$B = M \oplus N$. If$z = x + y$ is the unique expression of a vector z in B as the sum of vectors x and y in M and N, then a new norm can be defined on the linear space B by$\|z\|' = \|x\| + \|y\|$. Prove that this actually isa norm. If B' symbolizes the linear space B equipped with this new norm, prove that B' is a Banach space if M and N are closed in B.
3. Prove Eq. (9) for the case of an arbitrary positive integer n.

1.2 : CONTINUOUS LINEAR TRANSFORMATIONS

Let N and N' be normed linear spaces with the same scalars, and let T be a linear transformation of N into N'. When we say that T is continuous, we mean that it is continuous as a mapping of the metric space N into the metric space N'. By Theorem 13-B, this amounts to the condition that $x_n \to x$ in $N \Rightarrow T(x_n) \to T(x)$ in N'. Our main purposein this section is to convert the requirement of continuity into several more useful equivalent forms and to show that the set of all continuous linear transformations of N into N' can itself be made inte a normed linear space in a natural way.

Theorem 1.2 : Let N and N' be normed linear spaces and T a linear transformation of N into N'. Then the following conditions on T are all equivalent to one another:

(1) T is continuous;

(2) T is continuous at the origin, in the sense that $x_n \to 0 \Rightarrow T(x_n) \to 0$;

(3) there exists a real number $K > 0$ with the property that $\|T(x)\| \leq K\|x\|$ for every $x \in N$;

(4) if $S = \{x: \|x\| \leq 1\}$ ds the closed unit sphere in N, then its image $T(S)$ is a bounded set in N'.

Proof.

(1)$\Leftrightarrow$ (2). If T is continuous, then since $T(0) = 0$ it is certainly continuous at the origin. On the other hand, if T is continuous at the origin, then$x_n \to x \Leftrightarrow x_n - x \to 0 \Rightarrow T(x_n - x) \to 0 \Leftrightarrow T(x_n) - T(x) \to 0 \Leftrightarrow T(x_n) = T(x)$, so T is continuous.

(2) $\Leftrightarrow$ (3). It is obvious that (3) $\Rightarrow$ (2), for if such a K exists, then $x_n \to 0$ clearly implies that$T(x_n) \to 0$.

To show that (2) $\Rightarrow$ (3), we assume that there is no such K.

It follows from this that for each positive integer n we can find a vector x, such that $\|T(x_n)\| > n\|x_n\|$ or equivalently, such that $\left\|T\left(\frac{x_n}{n\|x_n\|}\right)\right\| > 1$.

If we now put

$$y_n = \frac{x_n}{n\|x_n\|},$$

then it is easy to see that $y_n \to 0$ but$T(y_n) \nrightarrow 0$, so T is not continuous at the origin.

(3) $\Leftrightarrow$ (4). Since a non-empty subset of a normed linear space is bounded $\Leftrightarrow$ it is contained in a closed sphere centered on the origin, it is evident that (3) $\Rightarrow$ (4); for if$\|x\| \leq 1$,

then$\|T(x)\| \le K$. To show that (4) ⇒ (3), we assume that $T(S)$ is contained in a closed sphere of radius K centered on the origin.

If $x = 0$, then $T(x) = 0$, and clearly $\|T(x)\| \le K\|x\|$; and if $x \ne 0$, then$\frac{x}{\|x\|} \in S$, and therefore$\left\|T\left(\frac{x}{\|x\|}\right)\right\| \le K$, so again we have$\|T(x)\| \le K\|x\|$.

If the linear transformation T in this theorem satisfies condition (3), so that there exists a real number $K > 0$ with the property that

$$\|T(x)\| \le K\|x\|$$

for every x, then K is called a bound for T, and such a T is often referred to as a bounded linear transformation.

According to our theorem, T is bounded ⇔ it is continuous, so these two adjectives can be used interchangeably.

We now assume that T is continuous, so that it satisfies condition (4), and we define its norm by

$$\|T\| = \sup\{\|T(x)\|: \|x\| \le 1\} \qquad (1)$$

When $N \ne \{0\}$, this formula can clearly be written in the equivalent form

$$\|T\| = \sup\{\|T(x)\|: \|x\| = 1\} \qquad (2)$$

It is apparent from the proof of Theorem A that the set of all bounds for T equals the set of all radii of closed spheres centered on the origin which contain $T(S)$. This yields yet another expression for the norm of T,

namely,

$$\|T\| = \inf\{K: K \ge 0 \; and \; \|T(x)\| \le K\|x\| \; for \; all \; x\}; \qquad (3)$$

and from this we see at once that

$$\|T(x)\| \leq \|T\|\|x\| \qquad (4)$$

for all x.

We now denote the set of all continuous (or bounded) linear transformations of N into N' by $\mathcal{C}(N, N')$, where the letter "$\mathcal{C}$" is intended to suggest the adjective ''bounded."' It is a routine matter to verify that this set is a linear space with respect to the pointwise linear operations defined by Eqs. (1) and (2) and to show that formula (1) actually does define a norm on this linear space. We summarize and extend these remarks in

Theorem 1.3 : If N and N' are normed linear spaces, then the set $\mathcal{C}(N, N')$ of all continuous linear transformations of N into N' is itself a normed linear space with respect to the pointwise linear operations and the norm defined by (1). Further, if N' is a Banach space, then $\mathcal{C}(N, N')$ is also a Banach space.

Proof.

We leave to the reader the simple task of showing that $\mathcal{C}(N, N')$ is a normed linear space, and we prove that this space is complete when N' is.

Let $\{T_n\}$ be a Cauchy sequence in $\mathcal{C}(N, N')$.

If x is an arbitrary vector in N, then $\|T_m(x) - T_n(x)\| = \|(T_m - T_n)(x)\| \leq \|T_m - T_n\|\|x\|$shows that $\{T_n(x)\}$ is a Cauchy sequence in N'; and since N' is complete, there exists a vector in N'-- we denote it by $T(z)$—such that $T_n(x) \to T(x)$.

This defines a mapping T of N into N', and by the joint con- tinuity of addition and scalar multiplication, T is easily seen to be a linear transformation.

To conclude the proof, we have only to show that T is continuous and that $T_n \to T$ with respect to the norm on $\mathcal{C}(N, N')$.. By the inequality (1), the norms of the terms of a Cauchy sequence in a normed linear space form a bounded set of numbers, so

$$\|T(x)\| = \|\lim T_n(x)\| = \lim\|T_n(x)\| \leq \sup(\|T_n\|\|x\|) = (\sup\|T_n\|)\|x\|$$

shows that T has a bound and is therefore continuous.

It remains to be proved that $\|T_n - T\| \to 0$. Let $\epsilon > 0$ be given, and let n_0 be a positive integer such that

$m, n \geq n_0 \Rightarrow \|T_m = T_n\| < \epsilon$. If$\|x\| \leq 1$and $m, n > n_0$, then

$$\|T_m(x) - T_n(x)\| = \|(T_m - T_n)(x)\|$$

$$= \|T_m - T_n\|\|x\|$$

$$\leq \|T_m - T_n\|$$

$$< \epsilon$$

We now hold m fixed and allow n to approach∞, and we see that $\|T_m(x) - T_n(x)\| \to \|T_m(x) - T(x)\|$, from which we conclude that $\|T_m(x) - T(x)\| < \epsilon$ for all $m > n_0$ and all x such that $\|x\| < 1$. This shows that $\|T_m - T\| \leq \epsilon$ for all $m > n_0$, and the proof is complete.

Let N be a normed linear space. We call a continuous linear transformation of N into itself an operator on N, and we denote the normed linear space of all operators on N by $\mathcal{C}(N)$ instead of $\mathcal{C}(N, N)$.

Theorem 1.3 shows that $\mathcal{C}(N)$ is a Banach space when N is. Furthermore, if operators are multiplied in accordance with formula (3), then $\mathcal{C}(N)$ is an algebra in which multiplication is related to the norm by

$$\|TT'\| \leq \|T\|\|T'\| \tag{5}$$

This relation is proved by the following computation:

$$\|TT'\| = \sup\{\|(TT')(x)\| : \|x\| \leq 1\}$$

$$= \sup\{\|T(T'(x))\| : \|x\| \leq 1\}$$

$$\leq \sup\{\|T\|\|T(x)\| : \|x\| \leq 1\}$$

$$= \|T\| \sup\{\|T'(x)\| : \|x\| \leq 1\}$$

$$= \|T\|\|T'\|$$

We know from the previous section that addition and scalar multiplication in $\mathcal{C}(N)$ are jointly continuous, as they are in any normed linear space. Property (5) permits us to conclude that multiplication is also jointly continuous:

$$T_n \to T \text{ and } T_n' \to T' \Rightarrow T_n T_n' \to TT'$$

This follows at once from

$$\|T_n T_n' - TT'\| = \|T_n(T_n' - T') + (T_n - T)T'\|$$

$$\leq \|T_n\|\|T_n' - T'\| + \|T_n - T\|\|T'\|$$

We also remark that when $N \neq \{0\}$, then the identity transformation I is an identity for the algebra $\mathcal{C}(N)$. In this case, we clearly have

$$\|I\| = 1 \qquad (6)$$

for $\|I\| = \sup\{\|I(x)\| : \|x\| \leq 1\} = \sup\{\|x\| : \|x\| \leq 1\} = 1$.

We complete this section with some definitions which will often be useful in our later work. Let N and N' be normed linear spaces. An isometric isomorphism of N into N' is a one-to-one linear transformation T of N into N' such that $\|T(x)\| = \|x\|$ for every x in N; and N is said to be zsometrically isomorphic to N' if there exists an isometric isomorphism of N onto N'. This terminology enables us to give precise meaning to the statement that one normed linear space is essentially the same as another.

Problems

1. If M is a closed linear subspace of a normed linear space N, and if T is the natural mapping of N onto $\frac{N}{M}$ defined by $T(z) = x + M$, show that T is a continuous linear transformation for which $\|T\| \leq 1$.

2. If T is a continuous linear transformation of a normed linear space N into a normed linear space N', and if M is its null space, show that T induces a natural linear transformation T' of $\frac{N}{M}$ into N' andthat $\|T'\| = \|T\|$.
3. Let N and N' be normed linear spaces with the same scalars. If N is infinite-dimensional and $N' \neq \{0\}$, show that there exists a linear transformation of N into N' which is not continuous. (We shall see in Problem 7 that if N is finite-dimensional, then every linear transformation of N into N' is automatically continuous.)
4. Let a linear space L be made into a normed linear space in two ways, and let the two norms of a vector x be denoted by $\|x\|$ and $\|x\|'$. These norms are said to be equivalent if they generate the sametopology on L. Show that this is the case $\Leftrightarrow$ there exist two positive real numbers K_1, and K_2 such that $K_1\|x\| \leq \|x\|' \leq K_2\|x\|$for all x. (If L is finite-dimensional, then any two norms defined on it are equivalent. See Problem 7.)
5. If n is a fixed positive integer, the spaces $l_p^n (1 \leq p < \infty)$ consist of a single underlying linear space with different norms defined on it. Show that these norms are all equivalent to one another. (Hint: show that convergence with respect to each norm amounts to coordinatewise convergence.)

1.3 : THE HAHN-BANACH THEOREM

One of the basic principles of strategy in the study of an abstract mathematical system can be stated as follows: consider the set of all structure-preserving mappings of that system into the simplest system of the same type. This principle is richly fruitful in the structure theory (or representation theory) of groups, rings, and algebras, and we shall see in the next section how it works for normed linear spaces.

We have remarked that the spaces R and C are the simplest of all normed linear spaces. If N is an arbitrary normed linear space, the above principle leads us to form the set of all continuous linear trans- formations of N into F or C, according as N is real or complex. This set— it is $\mathcal{C}(N, R)$ or $\mathcal{C}(N, C)$—is denoted by N^* and is called the conjugate space of N. The elements of N^* are called continuous linear functionals, or more briefly, functionals.It follows from our work in the previous section that if these functionals are added and multiplied by scalars pointwise, and if the norm of a functional f is defined by

$$\|f\| = \sup\{|f(x)|: \|x\| \leq 1\}$$

$$= \inf\{K: K \geq 0 \; and \; |f(x)| \leq K\|x\| \; for \; all \; x\},$$

then N^* is a Banach space.

When we consider various specific Banach spaces, the problem arises of determining the concrete nature of the functionals associated with these spaces. It is not our aim in this section to explore the ample body of theory which centers around this problem, and in any case, the machinery necessary for such an enterprise (mostly the theory of measure and integration) is not available to us. Nevertheless, for the reader who may have the required background, we mention some of the main facts with- out proof.

Let X be a measure space with measure m, and let p be a given real number such that $1 < p < \infty$. Consider the Banach space L_p, of all measurable functions f defined on X for which $|f(x)|^p$ is integrable. If g is a function in L_p, where $\frac{1}{p} + \frac{1}{q} = 1$, we define a function F_s, on L_p, by

$$F_g(f) = \int f(x)g(x)\, dm(x)$$

The Hélder inequality for integrals mentioned at the end of Problem 46-4 shows that

$$|F_g(f)| = \left|\int f(x)g(x)\, dm(x)\right|$$

$$\leq \int |f(x)g(x)dm(x)|$$

$$\leq \|f\|_g \|g\|_q$$

We conclude from this that F_g, is a well-defined scalar-valued linear function on L, with the property that$\|F_g\| \leq \|g\|_q$, and is therefore a functional on L_p. It can be shown that equality holds here, so that

$$\|F_g\| \leq \|g\|_q$$

It can also be shown that every functional on L_p, arises in this way, so the mapping $g \rightarrow F_p$, (which is clearly linear) is an isometric isomorphism of L_q onto L_p^*. This statement is usually expressed by writing

$$L_p^* = L_q \tag{1}$$

where the equality sign is to be interpreted in the sense just explained. If we specialize these considerations to n-tuples of scalars, we see that (1) becomes

$$\left(l_p^n\right)^* = l_q^n \tag{2}$$

Further, it can be shown that

$$(l_1^n)^* = l_\infty^n \tag{3}$$

and that

$$(l_\infty^n)^* = l_1^n \tag{4}$$

We sketch proofs of (2), (3), and (4) in the problems. When we consider sequences of scalars, then for $1 < p < \infty$ we have the following special case of (1):

$$l_p^* = l_q \tag{5}$$

If $p = 1$, we obtain a natural extension of (3):

$$l_1^* = l_\infty \tag{6}$$

The corresponding extension of (4) is another matter, for it is false that $l_\infty^* = l_1$. Instead, we have

$$c_0^* = l_1 \tag{7}$$

What is l_∞^*? We saw in Sec. 46 that l_∞ is a special case of $\mathcal{C}(X)$, so this question leads naturally to the problem of determining the nature of the conjugate space $\mathcal{C}^*(X)$.

The classic solution of this problem for a space X which is compact Hausdorff (or even normal) is known as the Riesz representation theorem, and it depends on some of the deeper parts of the theory of measure and integration (see Dunford and Schwartz [8, pp. 261-265]). The situation is somewhat simpler for the case in which X is an interval $[a, b]$ on the real line, but even here an adequate treatment requires a knowledge of Stieltjes integrals (see Riesz and Sz.-Nagy [35, secs. 49-51]).

Most of the theory of conjugate spaces rests on the Hahn-Banach theorem, which asserts that any functional defined on a linear subspace of a normed linear space can be extended linearly and continuously to the whole space without increasing its norm. The proof is rather complicated, so we begin with a lemma which serves to isolate its most difficult parts.

Lemma. Let M be a linear subspace of a normed linear space N, and let f be a functional defined on M. If x_0 is a vector not in M, and if

$$M_0 = M + [x_0]$$

is the linear subspace spanned by M and x_0, then f can be extended to a functional f_0 defined on M_0 such that $\|f_0\| = \|f\|$.

Proof.

We first prove the lemma under the assumption that N is a real normed linear space. We may assume, without loss of generality, that $\|f\| = 1$.

Since x_0 is not in M, each vector y in M_0 is uniquely expressible in the form $y = x + \alpha x_0$ with $x$in M.

It is clear that the definition $f_0(x + \alpha x_0) = f_0(x) + \alpha f_0(x_0) = f(x) + \alpha r_0$ extends f linearly to M_0, for every choice of the real number $r_0 = f_0(x_0)$.

Since we are trying to arrange matters so that $\|f_0\| = 1$, our problem is to show that $r_0$can be chosen in such a way that $|f_0(x + \alpha x_0)| \leq \|x + \alpha x_0\|$ for every x in M and every $\alpha \neq 0$. Since $f_0(x + \alpha x_0) = f(x) + \alpha r_0$, this inequality can be written as

$$-\|x + \alpha x_0\| \leq f(x) + \alpha r_0 \leq \|x + \alpha x_0\|$$

$$or \ -f(x) - \|x + \alpha x_0\| \leq \alpha r_0 \leq -f(x) + \|x + \alpha x_0\|$$

which in turn is equivalent to

$$-f\left(\frac{x}{\alpha}\right) - \left\|\frac{x}{\alpha} + x_0\right\| \leq r_0 \leq -f\left(\frac{x}{\alpha}\right) + \left\|\frac{x}{\alpha} + x_0\right\| \qquad (8)$$

We now observe that for any two vectors x_1 and x_2 in M we have

$$f(x_2) - f(x_1) = f(x_2 - x_1)$$

$$\leq |f(x_2 - x_1)|$$

$$\leq \|f\|\|x_2 - x_1\|$$

$$= \|x_2 - x_1\|$$

$$= \|(x_2 - x_0) - (x_1 + x_0)\|$$

$$= \|x_2 + x_0\| + \|x_1 + x_0\|$$

$$so, -f(x_1) - \|x_1 + x_0\| \leq -f(x_2) + \|x_2 + x_0\| \qquad (9)$$

If we define two real numbers a and b by

$$a = \sup\ \{-f(x) - \|x + x_0\|: x \in M\}$$

$$and \quad a = \inf\{-f(x) + \|x + x_0\|: x \in M\}$$

then (9) shows that $a \leq b$. If we now choose r_0 to be any real number such that $a \leq r_0 \leq b$, then the required inequality (8) is satisfied and this part of the proof is complete.

We next use the result of the above paragraph to prove the lemma for the case in which N is complex. Here f is a complex-valued functional defined on M for which $\|f\| = 1$.

We begin by remarking that a complex linear space can be regarded as a real linear space by simply restricting the scalars to be real numbers.

If g and h are the real and imaginary parts of f, so that $f(x) = g(x) + ih(x)$ for every x in M, then both g and h are easily seen to be real-valued functionals on the real space M; and since $\|f\| = 1$, we have $\|g\| \leq 1$.

The equation

$$f(ix) = if(x)$$

together with $f(ix) = g(ix) + ih(ix)$ and

$$if(x) = i(g(z) + ih(x)) = ig(x) - h(x)$$

shows that $h(x) = -g(ix)$, so we can write $f(x) = g(x) - ig(ix)$. By the above paragraph, we can extend g to a real-valued functional go on the real space My in such a way that $\|g_0\| = \|g\|$, and we define f_0 for x in M_0 by $f_0(x) = g_0(x) — ig_0(ix)$. It is easy to see that f_0, is an extension of f from M to M_0, that $f_0(x + y) = f_0(x) + f_0(y)$, and that $f_0(\alpha x) = \alpha f_0(x)$ for all real α's. The fact that the property last stated is also valid for all complex α's is a direct consequence of

$$f_0(ix) = g_0(ix) - ig_0(i^2 x) = g_0(ix) + ig_0(x) = i(g_0(x) - ig_0(ix)) = if_0(x),$$

so f_0 is linear as a complex-valued function defined on the complex space M_0. All that remains to be proved is that$\|f_0\| = 1$, and we dispose of this by showing that if x is a vector in M_0, for which $\|x\| = 1$, then $|f_0(x)| \leq 1$. If $f_0(x)$ is real, this follows from $f_0(x) = g_0(x)$ and$g_0 \leq 1$. If $f_0(x)$ is complex, then we can write $f_0(x) = re^{i\theta}$ with $r > 0$, so

$$|f_0(x)| = r = e^{-i\theta} f_0(x) = f_0\left(e^{-i\theta} x\right)$$

and our conclusion now follows from $\left\|e^{-i\theta} x\right\| = \|x\| = 1$ and the fact that $f_0\left(e^{-i\theta} x\right)$is real.

Theorem 1.4 (The Hahn-Banach Theorem) : Let M be a linear subspace of a normed linear space N, and let f be a functional defined on M. Then f can be extended to a functional f_0 defined on the whole space N such that $\|f_0\| = \|f\|$.

Proof.

The set of all extensions of f to functionals g with the same norm defined on subspaces which contain M is clearly a partially ordered set with respect to the following relation: $g_1 \leq g_2$ means that the domain of g_1: is contained in the domain of g_2, and $g_2(x) = g_1(x)$ for all x in the domain of g_1.

It is easy to see that the union of any chain of extensions is also an extension and is therefore an upper bound for the chain. Zorn's lemma now implies that there exists a maximal extension fy. We complete the proof by observing that the domain of f) must be the entire space N, for otherwise it could be extended further by our lemma and would not be maximal.

As we stated in the introduction to this chapter, the main force of the Hahn-Banach theorem lies in the guarantee it provides that any Banach space (or normed linear space) has a rich supply of functionals. This property is to be understood in the sense of the following two theorems, on which most of its applications depend.

Theorem 1.5 : If N is a normed linear space and x_0 is a non-zero vector in N, then there exists a functional f_0 in $N*$ such that $f_0(x_0) = \|x_0\|$ and $f_0 = 1$.

Proof.

Let $M = \{\alpha x_0\}$ be the linear subspace of N spanned by x_0, and define f on M by $f(\alpha x_0) = \alpha\|x_0\|$. It is clear that f is a functional on M such that $f(x_0) = \|x_0\|$ and $\|f\| = 1$. By the Hahn-Banach theorem, f can be extended to a functional f_0 in $N*$ with the required properties.

Among other things, this result shows that $N*$ separates the vectors in N, for if x and y are any two distinct vectors, so that $x - y \neq 0$, then there exists a functional f in N* such that $f(x - y) \neq 0$, or equivalently, $f(x) \neq f(y)$.

Theorem 1.6 : If M is a closed linear subspace of a normed linear space N and x_0 is a vector not in M, then there exists a functional f_0 in N* such that $f_0(M) = 0$ and $f_0(x_0) \neq 0$.

Proof.

The natural mapping T of N onto N/M (see Problem 47-1) is a continuous linear transformation such that $T(M) = 0$ and

$$T(x_0) = x_0 + M \neq 0$$

By Theorem 1.5, there exists a functional f in (N/M)* such that

$$f(x_0 + M) \neq 0$$

If we now define f_0 by $f_0(x) = f(T(x))$, then f_0 is easily seen to have the desired properties.

These theorems play a critical role in the ideas developed in the following sections, and their significance will emerge quite clearly in the proper context.

Problems

1. Let M be a closed linear subspace of a normed linear space N, and let x_0 be a vector not in M. If d is the distance from x to M, show that there exists a functional f_0 in N* such that$f_0(M) = 0, f_0(x_0) = 1$, and $\|f_0\| = 1/d$.
2. Prove that a normed linear space N is separable if its conjugate space N* is. (Hint: let $\{f_n\}$ be a countable dense set in N* and $\{x_n\}$ a corresponding set in N such that $\|x_n\| \leq 1$ and$|f_n(x_n)| \geq \|f_n\|/2$; let M be the set of all linear combinations of the x_n's whose coefficients are rational or—if N is complex—have rational real and imaginary parts; and use Theorem C to show that M = N.) We remark that N* need not be separable when N is, for 1, is easily proved to be separable, $l_1^* = l_\infty$ and l_∞ is not separable (see Problem 18-4).

1.4 : THE NATURAL IMBEDDING OF N IN N**

Since the conjugate space N* of a normed linear space N is itself a normed linear space, it is possible to form the conjugate space (N*)* of N*. We denote this space by N**, and we call it the second conjugate space of N.

The importance of N** rests on the fact that each vector x in N gives rise to a functional F_x, in N**. If we denote a typical element of N* by f, then F, is defined by

$$F_x(f) = f(x)$$

In other words, we invert the usual practice by regarding the symbol $f(x)$ as specifying a function of f for each fixed x, and we emphasize this point of view by writing $f(x)$ in the form $F_x(f)$. A simple manipulation of the definition shows that F_x, is linear:

$$F_x(\alpha f + \beta g) = (\alpha f + \beta g)(x)$$

$$= \alpha f(x) + \beta g(x)$$

$$= \alpha F_x(f) + \beta F_x(g)$$

If we now compute the norm of F_x, we see that

$$\|F_x\| = \sup\{|F_x(x)|: \|f\| \leq 1\}$$

$$= \sup\{|f(x): \|f\| \leq 1|\}$$

$$\leq \sup\{\|f\|\|x\|: \|f\| \leq 1\}$$

$$\leq \|x\|$$

that equality holds here, so for each x in N we have

$$\|F_x\| = \|x\|$$

It follows from these observations that $x \rightarrow F_x$. is a norm-preserving mapping of N into N**. F is called the functional on N* induced by the vector z, and we refer to functionals of this kind as induced functionals. We next point out that the mapping $x \to F_x$, is linear and is therefore anisometric isomorphism of N into N**. To verify this, we must show that $F_{x+y}(f) = (F_x + F_y)(f)$ and $F_{\alpha x}(f) = (\alpha F_x)(f)$ for every f in N*. The first of these relations follows from

$$F_{x+y}(f) = f(x+y)$$

$$= f(x) + f(y)$$

$$= F_x(f) + F_y(f)$$

$$= (F_x + F_y)(f)$$

and the second is proved similarly. The isometric isomorphism $x \to F_x$, is called the natural imbedding of N in N**, for it allows us to regard N as part of N** without altering any of its structure as a normed linear space.

We write

$$N \subseteq N^{**}$$

where this set inclusion is to be understood in the sense just explained.

A normed linear space N is said to be reflexive if N = N** The spaces l_p, (and L_p) for 1 < p <∞ are reflexive, for $l_p^* = l_p$, and

$$l_p^{**} = l_p^* = l_p$$

It follows from Problem 48-3 that the spaces l_p^n for $1 \le p \le \infty$ are also reflexive. Since N** is complete, N is necessarily complete if it is reflexive.

If N is complete, however, it is not necessarily reflexive, as we see from c_0* = l_1; and c_0** = l_1* = l_∞. If X is a compact Hausdorff space, it can be shown that $\mathcal{C}$(X) is reflexive ⇔ X is a finite set.

There is an interesting criterion for reflexivity, which depends on the concept of the weak topology on a normed linear space N. This is defined to be the weak topology on N generated by the functions in N* in the sense of Sec. 19; that is, it is the weakest topology on N with respect to which all the functions in V* remain continuous.

The criterion referred to is the following: if B isa Banach space, and if S = $\{x: \|x\| \leq 1\}$ is its closed unit sphere, then B is reflexive ⇔ S is compact in the weak topology.

This fact is something one should know about Banach spaces, but we shall have no need for it ourselves, so we state it without proof.

Far more important for our purposes is the weak* topology on N*, which is defined to be the weak topology on N* generated by all the induced functionals F, in N**.

This situation is rather complicated, so we shall try to make clear just what is going on.

First of all, N* (like N) is a normed linear space, and it therefore has a topology derived from its character as a metric space. This is called the strong topology.

N** is the set of all scalar-valued linear functions defined on N* which are continuous with respect to its strong topology.

The weak topology on N* (like the weak topology on N) is the weakest topology on N* with respect to which all the functions in N** are continuous, and clearly this is weaker than its strong topology.

So far, as we have indicated, these concepts apply equally to N and N*.

However, since N* is the conjugate space of N, the natural imbedding enables us to consider N as part of N**.

We now form the weakest topology on N* with respect to which all the functions in N—regarded as a subset of N**—remain continuous.

This is the weak* topology, and it is evidently weaker than the weak topology. The weak* topology can be given a more explicit description, in which its defining subbasic open sets are displayed. Consider a vector x in N and its induced functional F_x, in N**. The weak* topology on N* is the weakest topology under which all such F,'s are continuous. If fy is an arbitrary element in N*, and if $\epsilon > 0$ is given, then theset

$$S(x, f_0, \epsilon) = \{f : f \in N^* \ and \ |F_x(f) - F_x(f_0)| < \epsilon\}$$

$$= \{f : f \in N^* \ and \ |f(x) - f_0(x)| < \epsilon\}$$

is an open set (in fact, a neighborhood of f_0) in the weak* topology.

Furthermore, the class of all sets of this kind, for all x's, f_0's, and ϵ's, is the defining open subbase for the weak* topology. The finite intersections of these sets constitute an open base for this topology, and the open sets themselves are all unions of these finite intersections.

We remark at this point that N* is a Hausdorff space with respect to its weak* topology. This follows at once from the fact that if f and g are distinct functionals in N*, then there must exist a vector x in N such that $f(x) \neq g(x)$; for if we put $\epsilon = |f(x) — g(x)|/3$, then $S(x, f, \epsilon)$ and $S(x, g, \epsilon)$ are disjoint neighborhoods of f and g in the weak* topology.

Let us now consider the closed unit sphere S* in N*, that is, the set $S^* = \{f : f \in N^* \ and \ \|f\| < 1\}$.

It is an easy consequence of Problem 2 that S* is compact in the strong topology ⇔ N is finite-dimensional, so the strong compactness of S* is a very stringent condition. If N is complete, it follows from Problem 3 and our unproved criterion for reflexivity that S* is compact in the weak topology ⇔ J is reflexive, so the weak compactness of S* is still a fairly substantial restriction. We state these facts to emphasize that the situation is quite different with the weak* topology, for here S* is always compact.

Theorem 1.7 : If N is a normed linear space, then the closed unit sphere S* in N* is a compact Hausdorff space in the weak* topology.

Proof.

We already know that S* is a Hausdorff space in this topology, so we confine our attention to proving compactness. With each vector x in N we associate a compact space C_x, where C_x is the closed interval $[-\|x\|, \|x\|]$ or the closed dise $\{x: |x| \leq \|x\|\}$, according as N is real or complex.

By Tychonoff's theorem, the product C of all the C_x's is also a compact space. For each x, the values f(x) of all f's in S* lie in C_x.

This enables us to imbed S* in C by regarding each f in S* as identical with the array of all its values at the vectors x in N.

It is clear from the definitions of the topologies concerned that the weak* topology on S* equals its topology as a subspace of C'; and since C is compact, it suffices to show that S* is closed as a subspace of C.

We show that if g is in $\bar{S}$*, then g isin S*. If we consider g to be a function defined on the index set N, then since g is in C we have $|g(x)| \leq \|x\|$ for every x in N. It therefore suffices to show that g is linear as a function defined on N.

Let $\epsilon > 0$ be given, and let x and y be any two vectorsin N. Every basic neighborhood of g intersects S*, so there exists an f in S* such that $|g(x) - f(x)| < \epsilon/3$, $|g(y) - f(y)| < \epsilon/$ 3and $|g(x + y) - f(x + y)| < \epsilon/3$. Since $f$is linear, $f(z + y) - f(x) - f(y) = 0$, and we therefore have

$$|g(x+y) - g(x) - g(y)| = |[g(x+y) - f(x+y)]| - [g(x) - f(x)] - [g(y) - f(y)]$$

$$\leq |g(x+y) - f(x+y)| + |g(x) - f(x)| + |g(y) - f(y)|$$

$$< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}$$

$$= \epsilon$$

The fact that this inequality is true for every $\epsilon > 0$ now implies that $g(x + y) = g(x) + g(y)$. We can show in the same way that

$$g(\alpha x) = \alpha g(x)$$

for every scalar α, so g is linear and the theorem is proved.

We are now in a position to keep the promise made in the last paragraph of Sec. 46, for the following result is an obvious consequence of our preceding work,

Theorem 1.8 : Let N be a normed linear space, and let S* be the compact Hausdorff space obtained by imposing the weak* topology on the closed unit sphere in N*. Then the mapping $x \rightarrow F_x$, where $F_x(f) = f(x)$ for each f in S*, ts an isometric isomorphism of N into $\mathcal{C}$(S*). If N is a Banach space, this mapping is an isometric isomorphism of N onto a closed linear subspace of $\mathcal{C}$(S*).

This theorem shows, in effect, that the most general Banach space is essentially a closed linear subspace of $\mathcal{C}$(X), where X is a compact Hausdorff space.

The purpose of representation theorems in abstract mathematics is to reveal the structures of complex systems in terms of simpler ones, and from this point of view, Theorem 8 is satisfying to a degree.

It must be pointed out, however, that we know next to nothing about the closed linear subspaces of $\mathcal{C}$(X), though we know a good deal about $\mathcal{C}$(X) itself. Theorem 8 is therefore somewhat less revealing than appears at first glance. We shall see in Chaps. 13 and 14 that the corresponding representation theorem for Banach algebras is much more significant and useful.

Problems

1. Let X be a compact Hausdorff space, and justify the assertion that $\mathcal{C}$(X) is reflexive if X is finite.
2. If N is a finite-dimensional normed linear space of dimension n, show that N* also has dimension n. Use this to prove that N is reflexive.
3. If B is a Banach space, prove that B is reflexive ⇔ B* is reflexive.
4. Prove that if B is a reflexive Banach space, then its closed unit sphere S is weakly compact.
5. Show that a linear subspace of a normed linear space is closed ⇔ it is weakly closed.

UNIT - 2

2.1 : THE OPEN MAPPING THEOREM

In this section we have our first encounter with basic theorems which require that the spaces concerned be complete. The following rather technical lemma is the key to these theorems.

Lemma : If B and B' are Banach spaces, and if T is a continuous linear transformation of B onto B', then the image of each open sphere centered on the origin in B contains an open sphere centered on the origin in B'.

Proof.

We denote by S_r and S_r' the open spheres with radius r centered on the origin in B and B'. It is easy to see that

$$T(S_r) = T(rS_1) = rT(S_1)),$$

so it suffices to show that $T(S_1)$ contains some S_r'.

We begin by proving that $\overline{T(S_1)}$ contains some S_r'. Since T is onto, we see that $B' = \cup_{n=1}^{\infty} T(S_n)$. B' is complete, so Baire's theorem implies that some $\overline{T(S_{n_0})}$ has an intcrior point yo, which may be assumed to lie in $T(S_{n_0})$. The mapping $y \to y - y_0$ is a homeomorphism of B' onto itself, so $\overline{T(S_{n_0})} - y_0$ has the origin as an interior point.

Since y_0 is in $T(S_{n_0})$, we have $T(S_{n_0}) - y_0 \subseteq T(S_{2n_0})$; and from this we obtain $\overline{T(S_{n_1})} - y_0 = \overline{T(S_{n_0}) - y_0} \subseteq \overline{T(S_{2n_0})}$, Which shows that the origin is an interior point of $\overline{T(S_{2n_1})}$ Multiplication by any non-zero scalar is a homeomorphism of B' onto itself, so $\overline{T(S_{2n_1})} = \overline{2n_0 T(S_1)} = 2n_0\overline{T(S_1)}$; and it follows from this that the origin is also an interior point of $\overline{T(S_1)}$, so$S_0' \subseteq \overline{T(S_1)}$ for some positive number ϵ.

We conclude the proof by showing that $S'_\epsilon \subseteq T(S_3)$, which is clearly equivalent to $S'_{\epsilon/3} \subseteq T(S_1)$. Let y be a vector in B' such that $\|y\| < \epsilon$.

Since y is in $\overline{T(S_1)}$, there exists a vector x_1; in B such that $\|x_1\| < 1$ and $\|y - y_1\| < \epsilon/2$, where $y_1 = T(x_1)$.

We next observe that $S'_{\frac{\epsilon}{2}} \subseteq \overline{T\left(S_{\frac{1}{2}}\right)}$, so there exists a vector x_2 in B such that $\|x_2\| < \frac{1}{2}$ and$\|(y - y_1) - y_2\| < \epsilon/4$, where $y_2 = T(x_2)$.

Continuing in this way, we obtain a sequence $\{x_n\}$ in B such that $\|x_n\| < 1/2^{n-1}$ and$\|y - (y_1 + y_2 + \cdots + y_n)\| < \epsilon/2^n$, where $y_n = T(x_n)$. If weput

$$s_n = x_1 + x_2 + \cdots + x_n$$

then it follows from $\|x_n\| < 1/2^{n-1}$ that $\{s_n\}$ isa Cauchy sequence in B for which

$$\|s_n\| \leq \|x_1\| + \|x_2\| + \cdots + \|x_n\|$$

$$< 1 + \frac{1}{2} + \cdots + 1/2^{n-1}$$

$$< 2$$

B is complete, so there exists a vector x in B such that $s_n \to x$; and $\|x\| = \|\lim s_n\| = \lim\|s_n\| \leq 2 < 3$ shows that x is in S_3;. All that remains is to notice that the continuity of T yields

$$T(x) = T(\lim s_n) = \lim T(s_n) = \lim(y_1 + y_2 + \cdots + y_n) = y,$$

from which we see that y is in $T(S_3)$.

This makes our main theorem easy to prove.

Theorem 2.1 (The Open Mapping Theorem) : If B and B' are Banach spaces, and tf T is a continuous linear transformation of B onto B', then T is an open mapping.

Proof.

We must show that if G is an open set in B, then T(G) is also an open set in B'. If y is a point in T(G), it suffices to produce an open sphere centered on y and contained in T(G).

Let x be a point in G such that T(x) = y.

Since G is open, zis the center of an open sphere— which can be written in the form x + $S_r$contained in G.

Our lemma now implies that T(S_r) contains some S'_{r},. It is clear that y + S'_{r1}, is an open sphere centered on y, and the fact that it is contained in T(@) follows at once from $y + S'_{r_1} \subseteq y + T(S_r) = T(x) + T(S_r) \subseteq T(G)$

Most of the applications of the open mapping theorem depend more directly on the following special case, which we state separately for the sake of emphasis.

Theorem 2.2 : A one-to-one continuous linear transformation of one Banach space onto another is a homeomorphism. In particular, if a one-to-one linear transformation T of a Banach space onto itself is continuous, then its inverse T^{-1} is automatically continuous.

As our first application of Theorem 2.2, we give a geometric characterization of the projections on a Banach space. The reader will recall from Sec. 44 that a projection E on a linear space L is simply an idempotent $E^2 = E$ linear transformation of L into itself. He will also recall that projections on L can be described geometrically as follows:

(1) a projection E determines a pair of linear subspaces M and N such that L = M⊕N, where M = $\{E(x): x \in L\}$and N = $\{x: E(x) = 0\}$ are the range and null space of E;

(2) a pair of linear subspaces M and N such that L = M⊕N determines a projection E whose range and null space are M and N (if $z = x + y$ is the unique representation of a vector in Las a sum of vectors in M and N, then E is defined by E(z) = 2).

These facts show that the study of projections on L is equivalent to the study of pairs of linear subspaces which are disjoint and span L. In the theory of Banach spaces, however, more is required of a projection than mere linearity and idempotence.

A projection on a Banach space B is an idempotent operator on B; that is, it is a projection on B in the algebraic sense which is also continuous.

Our present task is to assess the effect of the additional requirement of continuity on the geometric descriptions given in (1) and (2) above. The analogue of (1) is easy.

Theorem 2.3. If P is a projection on a Banach space B, and if M and N are its range and null space, then M and N are closed linear subspaces of B such that B = M @ N.

Proof.

P is an algebraic projection, so (1) gives everything except the fact that M and N are closed.

The null space of any continuous linear transformation is closed, so N is obviously closed; and the fact that 1 is also closed is a consequence of

$$M = \{P(x): x \in B\}$$

$$= \{x: P(x) = x\}$$

$$= \{x: (I - P)(x) = 0\}$$

which exhibits M as the null space of the operator $I - P$.

The analogue of (2) is more difficult, for Theorem B is needed in itsproof.

Theorem 2.4 : Let B be a Banach space, and let M and N be closed linear subspaces of B such that B= M⊕N. If $z = x + y$ is the unique representation of a vector in B as a sum of vectors in M and N, then the mapping P defined by P(z) = x is a projection on B whose range and null space are M and N.

Proof.

Everything stated is clear from (2) except the fact that P is continuous, and this we prove as follows. By Problem 46-2, if B' denotes the linear space B equipped with the norm defined by

$$\|z\|' = \|x\| + \|y\|$$

then B' is a Banach space; and since $\|P(z)\| = \|x\| \leq \|x\| + \|y\| = \|z\|'$, P is clearly continuous as a mapping of B' into B.

It therefore suffices to prove that B' and B have the same topology. If 7 denotes the identity mapping of B' onto B, then

$$\|T(z)\| = \|z\|$$

$$= \|x + y\|$$

$$\leq \|x\| + \|y\|$$

$$= \|z\|'$$

shows that T is continuous as a one-to-one linear transformation of B' onto B. Theorem B now implies that T is a homeomorphism, and the proof is complete.

This theorem raises some interesting and significant questions. Let M be a closed linear subspace of a Banach space B.

As we remarked at the end of Sec. 44, there is always at least one algebraic projection defined on B whose range is M, and there may be a great many.

However, it might well happen that none of these are continuous, and that conse- quently none are projections in our present sense. In the light of our theorems, this is equivalent to saying that there might not exist any closed linear subspace N such that B = M⊕N.

What sorts of Banach spaces have the property that this awkward situation cannot occur? We shall see in the next chapter that a Hilbert space which is a special type of Banach space has this property. We shall also see that this property is closely linked to the satisfying geometric structure which sets Hilbert spaces apart from general Banach spaces.

We now turn to the closed graph theorem. Let B and B' be Banach spaces. If we define a metric on the product B×B' by

$$d\big((x_1, y_1), (x_2, y_2)\big) = \max\{\|x_1 - x_2\|, \|y_1 - y_2\|\},$$

then the resulting topology is easily seen to be the same as the product topology, and convergence with respect to this metric is equivalent to coordinatewise convergence. Now let T be a linear transformation of B into B'. We recall that the graph of T is that subset of B×B' which consists of all ordered pairs of the form (x,T(x)). Problem 26-6 shows that if 7 is continuous, then its graph is closed as a subset of B×B'. In the present context, the converse is also true.

Theorem 2.5 (The Closed Graph Theorem). If B and B' are Banach spaces, and if T is a linear transformation of B into B', then T is continuous > its graph is closed.

Proof.

In view of the above remarks, we may confine our attention to proving that T is continuous if its graph is closed. We denote by B the linear space B renormed by $\|x\|_1 = \|x\| + \|T(x)\|$. Since

$$\|T(x)\| \leq \|x\| + \|T(x)\| = \|x\|_1$$

T is continuous as a mapping of B_1 into B'. It therefore suffices to show that B and B_1 have the same topology.

The identity mapping of B_1 onto B is clearly continuous, for $\|x\| \leq \|x\| + \|T(x)\| = \|x\|$. If we can show that B_1 is complete, then Theorem 2.2 will guarantee that this mapping is a homeomorphism, and this will conclude the proof.

Let $\{x_n\}$ be a Cauchy sequence in B_1. It follows that $\{x_n\}$ and $\{T(x_n)\}$ are also Cauchy sequences in B and B'; and since both of these spaces are complete, there exist vectors x and y in

B and B' such that $\|x_n - x\| \to 0$ and $\|T(x_n) - y\| \to 0$. Our assumption that the graph of T is closed in $B \times B'$ implies that (x, y) lies on this graph, so $T(x) = y$. The completeness of B_1 now follows from

$$\|x_n - x\|_1 = \|x_n - x\| + \|T(x_n - x)\|$$

$$= \|x_n - x\| + \|T(x_n) - T(x)\|$$

$$= \|x_n - x\| + \|T(x_n) - y\| \to 0$$

The closed graph theorem has a number of interesting applications to problems in analysis, but since our concern here is mainly with matters of algebra and topology, we do not pause to illustrate its uses in this direction.

Problems

1. Let a Banach space B be made into a Banach space B' by means of a new norm, and show that the topologies generated by these norms are the same if either is stronger than the other.
2. In the text, we used Theorem B to prove the closed graph theorem. Show that Theorem B is a consequence of the closed graph theorem.
3. Let T be a linear transformation of a Banach space B into a Banach space B'. If $\{f_i\}$ is a set of functionals in B'* which separates the vectors in B', and if f_iT is continuous for each f_i, prove that Tis continuous.

2.2 : THE CONJUGATE OF AN OPERATOR

We shall see in this section that each operator T on a normed linear space N induces a corresponding operator, denoted by 7* and called the conjugate of TJ', on the conjugate space N*. Our first task is to define T*, and our second is to investigate the properties of the mapping T→T*. We base our discussion on the following theorem.

Theorem 2.6 (The Uniform Boundedness Theorem) : Let B be a Banach space and N a normed linear space. If $\{T_i\}$ is a non-empty set of continuous linear transformations of B into N

with the property that $\{T_i(x)\}$ is a bounded subset of N for each vector x in B, then $\{\|T_i\|\}$ ts a bounded set of numbers; that is, $\{T_i\}$ is bounded as a subset of $\mathcal{C}(B, N)$.

Proof.

For each positive integer n, the set

$$F_n = \{x: x \in B \; and \; \|T_i(x)\| \leq n \; for \; all \; i\}$$

is clearly a closed subset of B, and by our assumption we have

$$B = \cup_{n=1}^{\infty} F_n$$

Since B is complete, Baire's theorem shows that one of the F_n's, say F_{n_0}, has non-empty interior, and thus contains a closed sphere So with center x_0 and radius $r_0 > 0$.

This says, in effect, that each vector in every set $T_i(S_0)$ has norm less than or equal to n_0; and for the sake of brevity, we express this fact. by writing $\|T_i(S_0)\| \leq n_0$.

It is clear that $S_0 - x_0$ is the closed sphere with radius r_0 centered on the origin, so $(S_0 - x_0)/r_0$ is the closed unit sphere S.

Since x_0 is in S_0, it is evident that $\|T_i(S_0 - x_0)\| \leq 2n_0$. This yields $\|T_i(S)\| \leq 2n_0/r_0$, so $\|T_i\| \leq 2n_0/r_0$ for every i, and the proof is complete.

This theorem is often called the Banach-Steinhaus theorem, and it has several significant applications to analysis. See, for example, Zygmund [46, vol. 1, pp. 165-168] or G4l [11]. For the purposes we have in view, our main interest is in the following simple consequence of it.

Theorem 2.7 : A non-empty subset X of a normed linear space N is bounded $\Leftrightarrow f(X)$ ts a bounded set of numbers for each f in $N*$.

Proof.

Since $|f(x)| \leq \|f\|\|x\|$, it is obvious that if X is bounded, then $f(X)$ is also bounded for each f.

In order to prove the other half of the theorem, it is convenient to exhibit the vectors in X by writing $X = \{x_i\}$. We now use the natural imbedding to pass from X to the corresponding subset $\{F_{s_i}\}$ of N**.

Our assumption that $f(X) = \{f(x_i)\}$ is bounded for each f is clearly equivalent to the assumption that $\{F_{s_i(f)}\}$ is bounded for each f, and since N* is complete, Theorem 2.1 shows that $\{F_{s_i}\}$ is a bounded subset of N**. We know that the natural imbedding preserves norms, so X is evidently a bounded subset of N.

We now turn to the problem of defining the conjugate of an operator on a normed linear space N.

Let L be the linear space of all scalar-valued linear functions defined on N.

The conjugate space N* is clearly a linear subspace of L.

Let T be a linear transformation of N into itself which is not necessarily continuous. We use T to define a linear transformation T”’ of L into itself, as follows.

If f is in L, then $T'(f)$ is defined by

$$[T'(f)](x) = f(T(x)) \tag{1}$$

We leave it to the reader to verify that $T'(f)$ actually is linear as a function defined on N, and also that T' is linear as a mapping of L into itself.

The following natural question now presents itself. Under what circumstances does T' map N* into N*? This question has a simple and elegant answer: T'(N*) ⊆N* ⇔ T is continuous. If we keep Theorem B in mind, the proof of this statement is very easy; for if S is the closed unit sphere in N, then T is continuous ⇔ T(S) is bounded ⇔ f(T(S)) is bounded for each f in N* ⇔ $[T'(f)]$(S) is bounded for each fin N* ⇔ T’(f) is in N* for each f in N*.

We now assume that the linear transformation T is continuous and is therefore an operator on N. The preceding developments allow us to consider the restriction of T' to a mapping of N* into itself. We denote this restriction by T*, and we call it the conjugate of T. The action of T* is given by

$$[T^*(f)](x) = f(T(x)) \tag{2}$$

in which—in contrast to (1)—f is understood to be a functional on N, and not merely a scalar-valued linear function. T* is clearly linear, and the following computation shows that it is continuous:

$$\|T^*\| = \sup\{\|T^*(f)\|: \|f\| \leq 1\}$$

$$= \sup\{\|T^*(f)(x)\|: \|f\| \; and \; \|x\| \leq 1\}$$

$$= \sup\{\left|f\big(T(x)\big)\right|: \|f\| \; and \; \|x\| \leq 1\}$$

$$\leq \sup\{\|f\|\|T\|\|x\|: \|f\| \; and \; \|x\| \leq 1\}$$

$$\leq \|T\|$$

Since $\|T\| = \sup\{\|T(x)\|: \|x\| \leq 1\}$, we see at once from Theorem 48-B that equality holds here, that is, that

$$\|T^*\| = \|T\| \tag{3}$$

The mapping T → T* is thus a norm-preserving mapping of $\mathcal{C}$(N) into $\mathcal{C}$(N*).

We continue in this vein by observing that the mapping T→T* also has the following pleasant algebraic properties:

$$(\alpha T_1 + \beta T_2)^* = \alpha T_1^* + \beta T_2^* \tag{4}$$

$$(T_1 T_2)^* = T_2^* T_1^* \tag{5}$$

$$I^* = I \tag{6}$$

The proofs of these facts are easy consequences of the definitions. We illustrate the principles involved by proving (5). It must be shown that $(T_1T_2)^*(f) = (T_2^*T_1^*)(f)$ for each f in N*, and this means that

$$[(T_1T_2)^*(f)](x) = [(T_2^*T_1^*)(f)](x)$$

for each f in N* and each x in N. A simple computation now shows that

$$[(T_1T_2)^*(f)](x) = f\big((T_1T_2)(x)\big)$$

$$= f\left(T_1\big(T_2(x)\big)\right)$$

$$= [T_1^*(f)]\big(T_2(x)\big)$$

$$= [(T_2^*T_1^*)(f)](x)$$

It may be helpful to the reader to have the following summary of the results of this discussion.

Theorem 2.8 : If T is an operator on a normed linear space N, then its conjugate T* defined by Eq. (2) is an operator on N*, and the mapping T → T* is an isometric isomorphism of $\mathcal{C}$(N) into $\mathcal{C}$(N*) which reverses products and preserves the identity transformation.

The general significance of the ideas developed here can be under- stood only in the light of the theory of operators on Hilbert spaces. Some preliminary comments on these matters are given in the introduction to the next chapter.

Problems

1. Let B be a Banach space and N a normed linear space. If $\{T_n\}$ isa sequence in $\mathcal{C}$(B,N) such that T(x) = lim $T_n(x)$ exists for each x in B, prove that T is a continuous linear transformation.
2. Let T be an operator on a normed linear space N. If N is considered to be part of N** by means of the natural imbedding, show that T** is an extension of 7. Observe that if N is reflexive, then T** = T.

3. Let T be an operator on a Banach space B. Show that 7 has an inverse $T^{-1} \Leftrightarrow$ T* has an inverse $(T^{*})^{-1}$, and that in this case $(T^{*})^{-1} = (T^{-1})^{*}$.

2.3 : THE DEFINITION AND SOME SIMPLE PROPERTIES

The Banach spaces studied in the previous chapter are little more than linear spaces provided with a reasonable notion of the length of a vector. The main geometric concept missing in an abstract space of this type is that of the angle between two vectors. The theory of Hilbert spaces does not hinge on angles in general, but rather on some means of telling when two vectors are orthogonal.

In order to see how to introduce this concept, we begin by considering the three-dimensional Euclidean space R^3.

A vector in R^3 is of course an ordered triple $x = (x_1, x_2, x_3)$ of real numbers, and its norm is defined by

$$\|x\| = (|x_1|^2 + |x_2|^2 + |x_3|^2)^{\frac{1}{2}}$$

In elementary vector algebra, the inner product of z and another vector $y = (y_1, y_2, y_3)$ is defined by

$$(x, y) = x_1y_1 + x_2y_2 + x_3y_3$$

and this inner product is related to the norm by

$$(x, x) = \|x\|^2$$

We assume that the reader is familiar with the equation

$$(x, y) = \|x\|\|y\| \cos\theta$$

where θ is the angle between x and y, and also with the fact that x and y are orthogonal precisely when $(x, y) = 0$.

Most of these ideas can readily be adapted to the three-dimensional unitary space C^3. For any two vectors $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in this space, we define their inner product by

$$(x, y) = x_1\overline{y_1} + x_2\overline{y_2} + x_3\overline{y_3} \qquad (1)$$

Complex conjugates are introduced here to guarantee that the relation

$$(x, x) = \|x\|^2$$

remains true. It is clear that the inner product defined by (1) is linear as a function of x for each fixed y, and is also conjugate-symmetric, in the sense that $\overline{(x, y)} = (y, x)$.

In this case, it is no longer possible to think of (x, y) as representing the product of the norms of x and y and the cosine of the angle between them, for (x, y) is in general a complex number. Nevertheless, if the condition $(x, y) = 0$ is taken as the definition of orthogonality, then this concept is just as useful here as it is in the real case.

With these ideas as a background, we are now in a position to give our basic definition. A Hilbert space is a complex Banach space whose norm arises from an inner product, that is, in which there is defined a complex function (x, y) of vectors x and y with the following properties:

1. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$
2. $\overline{(x, y)} = (y, x)$
3. $(x, x) = \|x\|^2$

It is evident that the further relation

$$(x, \alpha y + \beta z) = \overline{\alpha}(x, y) + \overline{\beta}(x, z)$$

is a direct consequence of properties (1) and (2).

The reader may wonder why we restrict our attention to complex spaces. Why not consider real spaces as well? As a matter of fact, we could easily do so, and many writers adopt this approach. There are a few places in this chapter where complex scalars are necessary, but the theorems involved are not crucial, and we could get along with real scalars without too much difficulty. It is only in the complex case, however, that the theory of operators on a Hilbert space assumes a really satisfactory form. This will appear with particular clarity in the next chapter, where we make essential use of the fact that every polynomial equation of the nth degree with complex coefficients has exactly n complex roots (some of which, of course, may be repeated). For this and other reasons, we limit ourselves to the complex case throughout the rest of this book.

The following are the main examples of Hilbert spaces. In accord- ance with the above remarks, the scalars in each example are understood to be the complex numbers.

Example 1. The space l_2^n, with the inner product of two vectors .

$$x = (x_1, x_2, \dots, x_n) \; and \; y = (y_1, y_2, \dots, y_n)$$

defined by

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}$$

It is obvious that conditions (1) to (3) are satisfied.

Example 2. The space l_2, with the inner product of the vectors

$$x = (x_1, x_2, \dots, x_n, \dots) \; and \; y = (y_1, y_2, \dots, y_n, \dots)$$

defined by

$$(x, y) = \sum_{i=1}^{\infty} x_i \overline{y_i}$$

The fact that this series converges—and thus defines a complex number— for each x and y in l_2 is an easy consequence of Cauchy's inequality.

Example 3. The space L_2 associated with a measure space X with measure m, with the inner product of two functions f and g defined by

$$(f, g) = \int f(x)\overline{g(x)}dm(x)$$

This Hilbert space is of course not part of the official content of this book, but we mention it anyway in case the reader has some knowledge of these matters.

As our first theorem, we prove a fundamental relation known as the Schwarz inequality.

Theorem 2.9 : If x and y are any two vectors in a Hilbert space, then $|(x, y)| \leq \|x\|\|y\|$.

Proof.

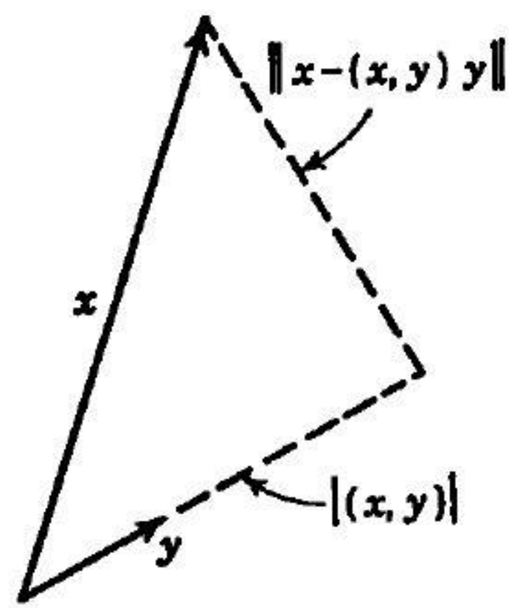

Fig. 2.**Schwar's Inequality**

When $y = 0$, the result is clear, for both sides vanish. When $y \neq 0$, the inequality is equivalent to $|(x, y/\|y\|)| \leq \|x\|$. We may therefore confine our attention to proving that if $\|y\| = 1$, then we have $|(x, y)| \leq \|x\|$ for all x. This is a direct consequence of the fact that

$$0 \leq \|x - (x,y)y\|^2$$

$$= (x - (x,y)y, x - (x,y)y)$$

$$= (x,x) - (x,y)\overline{(x,y)} - (x,y)\overline{(x,y)} + (x,y)\overline{(x,y)}(y,y)$$

$$= (x,x) - (x,y)\overline{(x,y)}$$

$$= \|x\|^2$$

$$= |(x,y)|^2$$

An inspection of Fig. 36 will reveal the geometric motivation for this computation.

It follows easily from Schwarz's inequality that the inner product in a Hilbert space is jointly continuous:

$$x_n \to x \; and \; y_n \to y \;\Rightarrow (x_n, y_n) \to (x,y)$$

To prove this, it suffices to observe that

$$|(x_n, y_n) - (x,y)| = |(x_n, y_n) - (x_n, y) + (x_n, y) - (x,y)|$$

$$\leq |(x_n, y_n) - (x_n, y)| + |(x_n, y) - (x,y)|$$

$$= |(x_n, y_n - y)| - |(x_n, x, y)|$$

$$\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|$$

A well-known theorem of elementary geometry states that the sum of the squares of the sides of a parallelogram equals the sum of the squares of its diagonals. This fact has an analogue in the present context, for in any Hilbert space the so-called parallelogram law holds:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

This is readily proved by writing out the expression on the left in terms of inner products:

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

$$=(x,x)+(x,y)+(y,x)+(y,y)+(x,x)-(x,y)-(y,x)+(y,y)$$

$$=2(x,x)+2(y,y)$$

$$=2\|x\|^2+2\|y\|^2$$

The parallelogram law has the following important consequence for our work in the next section.

Theorem 2.10 : A closed convex subset C of a Hilbert space H contains a unique vector of smallest norm.

Proof.

We recall from the definition in Problem 32-5 that since C is convex, it is non-empty and contains $(x + y)/2$ whenever it contains x and y.

Let $d = inf\{\|x\|: x \in C\}$. There clearly exists a sequence $\{x_n\}$ of vectors in C such that $\|x_n\| \to d$.

By the convexity of C, $(x_m + x_n)/2$ is in C and $\|(x_m + x_n)/2\| \geq d$, so $\|(x_m + x_n)\| \geq 2d$. Using the parallelogram law, we obtain

$$\|x_m+x_n\|^2=2\|x_m\|^2+2\|x_n\|^2-\|x_m+x_n\|^2$$

$$\leq 2\|x_m\|^2+2\|x_n\|^2-4d^2$$

and since $2\|x_m\|^2+2\|x_n\|^2-4d^2 \to 2d^2+2d^2-4d^2=0$, it follows that $\{x_n\}$ is a Cauchy sequence in C.

Since H is complete and C is closed, C is complete, and there exists a vector x in C such that $x_n \to x$. It is clear by the fact that $\|x\| = \|\lim x_n\| = \lim\|x_n\| = d$ that x is a vector in C with smallest norm.

To see that x is unique, suppose that x' is a vector in C other than x which also has norm d. Then $(x + x')/2$ is also in C, and another application of the parallelogram law yields

$$\left\|\frac{x + x'}{2}\right\| = \frac{\|x\|^2}{2} + \frac{\|x'\|^2}{2} - \left\|\frac{x - x'}{2}\right\|^2$$

$$< \frac{\|x\|^2}{2} + \frac{\|x'\|^2}{2}$$

$$= d^2$$

which contradicts the definition of d.

The parallelogram law has another interesting application, which depends on the fact that in any Hilbert space the inner product is related to the norm by the following identity:

$$4(x, y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 \qquad (2)$$

This is easily verified by converting the expression on the right into inner products.

Theorem 2.11 : If $B$is a complex Banach space whose norm obeys the parallelogram law, and if an inner product is defined on B by (2), then B is a Hilbert space.

Proof.

All that is necessary is to make sure that the inner product defined by (2) has the three properties required by the definition of a Hilbert space.

This is easy in the case of properties (2) and (3). Property (1) is best treated by splitting it into two parts:

$$(x + y, z) = (x, z) + (y, z)$$

and $(\alpha x, y) = \alpha(x, y)$. The first requires the parallelogram law, and the second follows from the first. We ask the reader (in Problem 6) to work out the details.

This result has no implications at all for our future work. However, it does provide a satisfying geometric insight into the place Hilbert spaces occupy among all complex Banach spaces: they are precisely those in which the parallelogram law is true.

Problems

1. Show that the series which defines the inner product in Example 2 is convergent.
2. The Hilbert cube is the subset of l_2, consisting of all sequences
$$x = \{x_1, x_2, \dots, x_n, \dots\}$$
such that $|x_n| \leq 1/n$ for all n. Show that this set is compact as a subspace of l_2
3. For the special Hilbert space l_2^n, use Cauchy's inequality to prove Schwarz's inequality.
4. Show that the parallelogram law is not true in l_2^n $(n > 1)$.
5. In a Hilbert space, show that if$\|x\| = \|y\| = 1$, and if $\epsilon > 0$ is given, then there exists $\delta > 0$ such that $\|(x + y)/2\| > 1 - \delta \Rightarrow \|x - y\| < \epsilon$. A Banach space with this property is said to be uniformly convex. See Taylor (41, p. 231].
6. Give a detailed proof of Theorem C.

2.4 : ORTHOGONAL COMPLEMENTS

Two vectors x and y in a Hilbert space H are said to be orthogonal (written $x \perp y$) if $(x, y) = 0$. Thesymbol $\perp$ is often pronounced "perp." Since $\overline{(x, y)} = (y, x)$, wehave$x \perp y \Leftrightarrow y \perp x$. Itis also clear that $x \perp 0$ for every x, and $(x, x) = \|x\|^2$ shows that 0 is the only vector orthogonal to itself. One of the simplest geometric facts about orthogonal vectors is the Pythagorean theorem:

$$x \perp y \Rightarrow \|x + y\|^2 = \|x - y\|^2 = \|x\|^2 + \|y\|^2$$

A vector x is said to be orthogonal to a non-empty set S (written $x \perp$S) if $x \perp y$ for every y in S, and the orthogonal complement of S—denoted by $S^\perp$—is the set of all vectors orthogonal to S. The following state- ments are easy consequences of the definition:

$$\{0\}^\perp = H; H^\perp = \{0\}$$

$$S \cap S^\perp \subseteq \{0\}$$

$$S_1 \subseteq S_2 \Rightarrow S_1^{\perp} \supseteq S_2^{\perp}$$

$S^{\perp}$ is a closed linear subspace of H.

It is customary to write $(S^{\perp})^{\perp}$ in the form $S^{\perp\perp}$. Clearly, $S \subseteq S^{\perp\perp}$

Let M be a closed linear subspace of H. We know that $M^{\perp}$ is also a closed linear subspace, and that M and M+ are disjoint in the sense that they have only the zero vector in common. Our aim in this section is to prove that $H = M \oplus M^{\perp}$, and each of our theorems is a step in this direction.

Theorem 2.12 : Let M be a closed linear subspace of a Hilbert space H, let x be a vector not in M, and let d be the distance from x to M. Then there exists a unique vector y_0 in M such that $\|x - y_0\| = d$.

Proof.

The set $C = x + M$ isa closed convex set, and d is the distance from the origin to C (see Fig. 2).

By Theorem, there exists a unique vector z_0 in C such that $\|z_0\| = d$. The vector $y_0 = x - z_0$is easily seen to be in M, and $\|x - y_0\| = \|z_0\| = d$.

The uniqueness of y_0 follows from the fact that if y_1 is a vector in M such that$y_1 \neq y_0$ and $\|x - y_1\| = d$, then $z_1 = x - y_1$ is a vector in C such that $z_1 \neq z_0$ and $\|z_1\| = d$, which contradicts the uniqueness of z_0.

We use this result to prove

Theorem 2.13 : If M is a proper closed linear subspace of a Hilbert space H, then there exists a non-zero vector z_0 in H such that $z_0 \perp M$.

Proof.

Let x be a vector not in M, and let d be the distance from x to M. By Theorem, there exists a vector y_0 in M such that $\|x - y_0\| = d$.

We define z_0 by $z_0 = x - y_0$ (see Fig. 37), and we observe that since $d > 0, z_0$ is a non-zero vector. We conclude the proof by showing that if y is an arbitrary vector in M, then $z_0 \perp y$. For any scalar α, we have

$$\|z_0 - \alpha y\| = \|x - (y_0 + \alpha y)\| \geq d = \|z_0\|$$

$$so\ \|z_0 - \alpha y\|^2 - \|z_0\|^2 \geq 0$$

$$and\ -\overline{\alpha}(z_0, y) - \alpha\overline{(z_0, y)} + |\alpha|^2\|y\|^2 \geq 0 \qquad (1)$$

If we put $\alpha = \beta(z_0, y)$ for an arbitrary real number β, then (1) becomes

$$-2\beta|(z_0, y)|^2 + \beta^2|(z_0, y)|^2\|y\|^2 \geq 0$$

If we now put $\alpha = |(z_0, y)|^2$ and $\|y\|^2$, we obtain

$$-2\beta a + \beta^2 ab \geq 0$$

$$so\ \ \beta a(\beta b - 2) \geq 0 \qquad (2)$$

for all real β. However, if $a > 0$, then (2) is obviously false for al sufficiently small positive β. We see from this that $a = 0$, which means that $z_0 \perp y$.

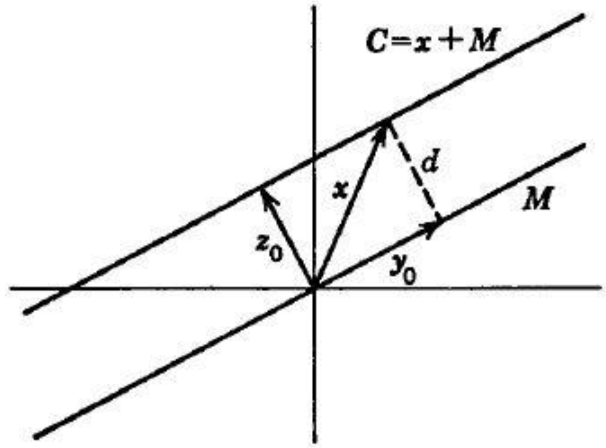

Fig. 37

This proof of Theorem B may strike the reader as being excessively dependent on ingenious computations. If so, he will be pleased to learn that the ideas developed in the next sec- tion can be used to provide another proof which is free of computation.

In order to state our next theorem, we need the following additional concept. Two non-empty subsets S; and S, of a Hilbert space are said to be orthogonal (written $S_1 \perp S_2$) if $x \perp y$ for all z in S_1 and y in S_2.

Theorem 2.14 : If M and N are closed linear subspaces of a Hilbert space H such that $M \perp N$, then the linear subspace $M + N$ is also closed.

Proof.

Let z be a limit point of $M + N$. It suffices to show that z is in $M + N$. There certainly exists a sequence $\{z_n\}$ in $M + N$ such that $z_n \rightarrow z$.

By the assumption that $M \perp N$, we see that M and N are disjoint, so each z_n, can be written uniquely in the form $z_n = x_n + y_n$, where x_n, isin M and y_n isin N.

The Pythagorean theorem shows that $\|z_m - z_n\|^2 = \|x_m - x_n\|^2 + \|y_m - y_n\|^2$so $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in $M$and N.

M and N are closed, and therefore complete, so there exist vectors x and y in M and N such that $x_n \to x$ and $y_n \to y$. Since $x + y$ is in $M + N$, our conclusion follows from the fact that $z = lim\, z_n = lim\,(x_n + y_n) = \lim x_n + \lim y_n = x + y$.

The way is now clear for the proof of our principal theorem.

Theorem 2.15 : If M is a closed linear subspace of a Hilbert space H, then $H = M \oplus M^{\perp}$.

Proof.

Since M and $M^{\perp}$ are orthogonal closed linear subspaces of H, Theorem C shows that $M + M^{\perp}$ is also a closed linear subspace of H.

We prove that $M + M^{\perp}$ equals H. If this is not so, then by Theorem B there exists a vector $z_0 \neq 0$ such that $z_0 \perp (M + M^{\perp})$.

This non-zero vector must evidently lie in $M^{\perp} \cap M^{\perp\perp}$; and since this is impossible, we infer that $H = M + M^{\perp}$.

To conclude the proof, it suffices to observe that since M and $M^{\perp}$ are disjoint, the statement that $H = M + M^{\perp}$ can be strengthened to $H = M \oplus M^{\perp}$.

The main effect of this theorem is to guarantee that a Hilbert space is always rich in projections.

In fact, if M is an arbitrary closed linear subspace of a Hilbert space H, then it shows that there exists a projection defined on H whose range is M and whose null space is M+. This satisfactory state of affairs is to be contrasted with the situation in a general Banach space.

Problems

1. If $S$is a non-empty subset of a Hilbert space, show that $S^{\perp} = S^{\perp\perp\perp}$.
2. If M is a linear subspace of a Hilbert space, show that M is closed $\Leftrightarrow M = M^{\perp\perp}$.
3. If S is a non-empty subset of a Hilbert space H, show that the set of all linear combinations of vectors in S is dense in $H \Leftrightarrow S^{\perp} = \{0\}$.

4. If Sisa non-empty subset of a Hilbert space H, show that $S^{\perp\perp}$ isthe closure of the set of all linear combinations of vectors in S. This is usually expressed by saying that $S^{\perp\perp}$ is the smallest closed linear subspace of H which contains S.

2.5 : ORTHONORMAL SETS

An orthonormal set in a Hilbert space H is a non-empty subset of H which consists of mutually orthogonal unit vectors; that is, it is a non-empty subset $\{e_i\}$ of H with the following properties:

1. $i \neq j \Rightarrow e_i \perp e_j$;
2. $\|e_i\| = 1$ for every i.

If H contains only the zero vector, then it has no orthonormal sets. If H contains a non-zero vector x, and if we normalize x by considering $e = x/\|x\|$, then the single-element set $\{e\}$ is clearly an orthonormal set. More generally, if $\{x_i\}$ is a non-empty set of mutually orthogonal non- zero vectors in H, and if the x_i's are normalized by replacing each of them by $e_i = x_i/\|x_i\|$, then the resulting set $\{e_i\}$ is an orthonormal set.

Example 1. The subset $\{e_1, e_2, \dots, e_n\}$ of l_2^n, where e_i is the n-tuple with 1 in the ithplace and 0's elsewhere, is evidently an orthonormal set in this space.

Example 2. Similarly, if e_n is the sequence with 1 in the nth place and 0's elsewhere, then $\{e_1, e_2, \dots, e_n, \dots\}$ isan orthonormal set in l_2.

At the end of this section, we give some additional examples taken from the field of analysis.

Every aspect of the theory of orthonormal sets depends in one way or another on our first theorem.

Theorem 2.16 : Let $\{e_1, e_2, \dots, e_n\}$ be a finite orthonormal set in a Hilbert space H. If x is any vector in H, then

$$\sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2 \tag{1}$$

$$further, \quad x - \sum_{i=1}^{n} (x, e_i) e_i \perp e_j \tag{2}$$

for each j.

Proof.

The inequality (1) follows from a computation similar to that used in proving Schwarz's inequality:

$$0 \leq \left\| x - \sum_{i=1}^{n} (x, e_i) e_i \right\|^2$$

$$= \left(x - \sum_{i=1}^{n} (x, e_i) e_i \, x - \sum_{j=1}^{n} (x, e_j) e_j \right)$$

$$= (x, x) - \sum_{i=1}^{n} (x, e_i)\overline{(x, e_i)} - \sum_{j=1}^{n} (x, e_j)\overline{(x, e_j)} + \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)\overline{(x, e_j)}\,(e_i, e_j)$$

$$= \|x\|^2 - \sum_{i=1}^{n} |(x, e_i)|^2$$

To conclude the proof, we observe that

$$\left(x - \sum_{i=1}^{n} (x, e_i) e_i \, e_j \right) = (x, e_j) - \sum_{i=1}^{n} (x, e_i)(e_i \, e_j) = (x, e_j) - (x, e_j) = 0$$

from which statement (2) follows at once.

The reader should note that the inequality (1) can be given the following loose but illuminating geometric interpretation: the sum of the squares of the components of a vector in various perpendicular directions does not exceed the square of the length of the vector itself.

This is usually called Bessel's inequality, though, as we shall see below, it is only a special case of a more general inequality with the same name.

In a similar vein, relation (2) says that if we subtract from a vector its com- ponents in several perpendicular directions, then the result has no com- ponent left in any of these directions.

Our next task is to prove that both parts of Theorem 2.16 generalize to the case of an arbitrary orthonormal set. The main problem hereis to show that the sums in (1) and (2) can be defined in a reasonable way when no restriction is placed on the number of e_i's under consideration. The key to this problem lies in the following theorem.

Theorem 2.17 : If $\{e_i\}$is an orthonormal set in a Hilbert space H, and if x is any vector in H, then the set $S = \{e_i: (x, e_i) \neq 0\}$is either empty or countable.

Proof.

For each positive integer n, consider the set

$$S_n = \{e_i: |(x, e_i)|^2 > \|x\|^2/n\}$$

By Bessel's inequality, S_n contains at most $n - 1$ vectors. The con- clusion now follows from the fact that $S = \cup_{n=1}^{\infty} S_n$

As our first application of this result, we prove the general form of Bessel's inequality.

Theorem 2.18 (Bessel's Inequality) : If $\{e_i\}$is an orthonormal set in a Hilbert space H, then

$$\sum |(x, e_i)|^2 \leq \|x\|^2 \qquad (3)$$

for every vector x in H.

Proof.

Our basic obligation here is to explain what is meant by the sum on the left of (3). Once this is clearly understood, the proof is easy.

As in the preceding theorem, we write $S = \{e_i : (x, e_i) \neq 0\}$. If S is empty, we define $\sum |(x, e_i)|^2$to be the number 0; and in this case, (3) is obviously true.

We now assume that S is non-empty, and we see from Theorem 2.17 that it must be finite or countably infinite.

If S is finite, it can be written in the form $S = \{e_1, e_2, \ldots, e_n, ..\}$for some positive integer n. In this case, we define $\sum |(x, e_i)|^2$ to be $\sum_{i=1}^{\infty} |(x, e_i)|^2$, which is clearly independent of the order in which the elements of S are arranged.

The inequality (3) now reduces to (1), which has already been proved. All that remains is to consider the case in which S is countably infinite.

Let the vectors in S be arranged in a definite order:

$$S = \{e_1, e_2, \ldots, e_n, ..\}$$

By the theory of absolutely convergent series, if $\sum_{i=1}^{\infty} |(x, e_i)|^2$ converges, then every series obtained from this by rearranging its terms also con- verges, and all such series have the same sum.

We therefore define $\sum |(x, e_i)|^2$ to be $\sum_{n=1}^{\infty} |(x, e_n)|^2$, and it follows from the above remark that $\sum |(x, e_i)|^2$ is a non-negative extended real number which depends only on S, and not on the arrangement of its vectors.

We conclude the proof by observing that in this case, (3) reduces to the assertion that

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq \|x\|^2 \qquad (4)$$

and since it follows from (1) that no partial sum of the series on the left of (4) can exceed $\|x\|^2$, it is clear that (4) itself is true.

The second part of Theorem A is generalized in essentially the same way.

Theorem 2.19 : If $\{e_i\}$ is an orthonormal set in a Hilbert space H, and if x is an arbitrary vector in H, then

$$x - \sum(x, e_i)e_i \perp e_j \tag{5}$$

for each j.

Proof.

As in the above proof, we define $\sum(x, e_i)e_i$ for each of the various cases, and we prove (5) as we go along. We again write

$$S = \{e_i : (x, e_i) \neq 0\}$$

When S is empty, we define $\sum(x, e_i)e_i$ to be the vector 0, and we observe that (5) reduces to the statement that $x - 0 = x$is orthogonal to each e_j, which is precisely what is meant by saying that S is empty. When S is non-empty and finite, and can be written in the form

$$S = \{e_1, e_2, \dots, e_n\}$$

we define $\sum(x, e_i)e_i$ to be $\sum_{i=1}^{n}(x, e_i)e_i$ and in this case, (5) reduces to (2), which has already been proved.

We may assume for the remainder of the proof that S is countably infinite. Let the vectors in S be listed in a definite order: $S = \{e_1, e_2, \dots, e_n, \dots\}$. We put $s_n = \sum_{i=1}^{n}(x, e_i)e_i$, and we note that for $m > n$ we have

$$\|s_m - s_n\|^2 = \left\| \sum_{i=n+1}^{m} (x, e_i)e_i \right\|^2 = \sum_{i=n+1}^{m} |(x, e_i)e_i|^2$$

Bessel's inequality shows that the series $\sum_{i=1}^{n}|(x,e_i)e_i|^2$ converges, so $\{s_n\}$ is a Cauchy sequence in H; and since H is complete, this sequence converges to a vector s, which we write in the form $s = \sum_{n=1}^{\infty}(x,e_n)e_n$.

We now define $\sum(x,e_i)e_i$ to be $\sum_{n=1}^{\infty}(x,e_n)e_n$, and deferring for a moment the question of what happens when the vectors in S are rearranged we observe that (5) follows from (2) and the continuity of the inner product:

$$\left(x - \sum(x,e_i)e_i\,,\mathrm{e_j}\right) = \left(x - s, e_j\right)$$

$$= \left(x,e_j\right) - \left(s,e_j\right)$$

$$= \left(x,e_j\right) - \left(\lim s_n\,, e_j\right)$$

$$= \left(x,e_j\right) - \lim(s_n\,,e_j)$$

$$= \left(x,e_j\right) - \left(x,e_j\right)$$

$$= 0$$

All that remains is to show that this definition of $\sum\left(x,e_j\right)e_i$ is valid, in the sense that it does not depend on the arrangement of the vectors in S.Let the vectors in S be rearranged in any manner:

$$S = \{f_1, f_2, \dots, f_n, \dots\}$$

We put $S_n' = \sum_{i=1}^{n}(x_i f_i)f_i$ and we see—as above—that the sequence $\{s_n'\}$ converges to a limit s', which we write in the forms$' = \sum_{n=1}^{\infty}(x_n f_n)f_n$.

We conclude the proof by showing that s' equals s. Let $\epsilon >\ 0$ be given, and let no be a positive integer so large that if $n \geq n_0$, then $\|s_n - s\| < \epsilon, \|s_n' - s'\| < \epsilon$ and $\sum_{i=n+1}^{\infty}|(x,e_i)|^2 < \epsilon^2$. For some positive integer $m_0 > n_0$, all terms of s_{n_0} occur among those of s_{m_1}' so $s_{m_1}' - s_{n_0}$is a finite sum of terms of the form $\left(x,e_j\right)e_i$for$i = n_0 + 1, n_0 + 2, \dots$.

This yields $\left\|s'_{m_0} - s_{n_0}\right\|^2 \leq \sum_{i=m_0+1}^{\infty} |(x, e_i)|^2 < \epsilon^2$ so $\left\|s_{m_0} - s_{n_0}\right\| < \epsilon$ and

$$\|s' - s\| \leq \left\|s' - s'_{m_0}\right\| + \left\|s'_{m_0} - s_{n_0}\right\| + \left\|s_{n_0} - s\right\|$$

$$< \epsilon + \epsilon + \epsilon$$

$$= 3\epsilon$$

Since ϵ is arbitrary, this shows that $s' = s$.

Let H be a non-zero Hilbert space, so that the class of all its orthonormal sets is non-empty.

This class is clearly a partially ordered set with respect to set inclusion. An orthonormal set $\{e_i\}$ in H is said to be complete if it is maximal in this partially ordered set, that is, if it is impossible to adjoin a vector e to $\{e_i\}$ in such a way that $\{e_i e\}$ is an orthonormal set which properly contains $\{e_i\}$.

Theorem 2.20 : Every non-zero Hilbert space contains a complete orthonormal set.

Proof : The statement follows at once from Zorn's lemma, since the union of any chain of orthonormal sets is clearly an upper bound for the chain in the partially ordered set of all orthonormal sets.

Orthonormal sets are truly interesting only when they are complete. The reasons for this are presented in our next theorem.

Theorem 2.21 : Let H be a Hilbert space, and let $\{e_i\}$ be an orthonormal set in H. Then the following conditions are all equivalent to one another:

(1) $\{e_i\}$is complete;

(2) $x \perp \{e_i\} \Longrightarrow x = 0$

(3) if $x$is an arbitrary vector in H, then $x = \sum (x, e_i) e_i$

(4) if $x$is an arbitrary vector in H, then $\|x\|^2 = \sum |(x, e_i)|^2$.

Proof.

We prove that each of the conditions (1), (2), and (3) implies the one following it and that (4) implies (1).

(1) ⇒ (2). If (2) is not true, there exists a vector $x \neq 0$ such that$x \perp \{e_i\}$. We now define e by $e = x/\|x\|$, and we observe that $\{e_i e\}$is an orthonormal set which properly contains $\{e_i\}$. This contradicts the completeness of $\{e_i\}$.

(2) ⇒ (3). By Theorem D, $x - \sum(x, e_i)e_i$ is orthogonal to $\{e_i\}$, so (2) implies that $x - \sum(x, e_i)e_i = 0$, or equivalently, that $x = \sum(x, e_i)e_i$.

(3) ⇒ (4). By the joint continuity of the inner product, the expression in (3) yields

$$\|x\|^2 = (x, x)$$

$$= \left(\sum(x, e_i)e_i, \sum\left(x, e_j\right)e_j\right)$$

$$= \sum(x, e_i)\overline{(x, e_i)}$$

$$= \sum|(x, e_i)|^2$$

(4) ⇒ (1). If $\{e_i\}$ is not complete, it is a proper subset of an orthonormal set $\{e_i e\}$. Since e is orthogonal to all the e_i's, (4) yields $\|e\|^2 = \sum|(e, e_i)|^2 = 0$, and this contradicts the fact that e is a unit vector.

There is some standard terminology which is often used in connection with this theorem. Let $\{e_i\}$ be a complete orthonormal set in a Hilbert space H, and let x be an arbitrary vector in H. The numbers (x, e_i) are called the Fourier coefficients of x, the expression $x = \sum(x, e_i)e_i$; is called the Fourier expansion of x, and the equation $\|x\|^2 = \sum|(x, e_i)|^2$ is called Parseval's equation—all with respect to the particular complete ortho- normal set $\{e_i\}$ under consideration. These terms come from the classical theory of Fourier series, as indicated in our next example.

Example 3. Consider the Hilbert space L_2 associated with the measure space $[0,2\pi]$, where measure is Lebesgue measure and integrals are Lebesgue integrals! This space essentially consists of all complex functions f defined on $[0,2\pi]$ which are Lebesgue measurable and squareintegrable, in the sense that

$$\int_0^{2\pi} |f(x)|^2\, dx < \infty$$

Its norm and inner product are defined by

$$\|f\| = \left(\int_0^{2\pi} |f(x)|^2\, dx\right)^{1/4}$$

$$and \quad (f,g) = \int_0^{2\pi} f(x)\,\overline{g(x)} dx$$

A simple computation shows that the functions e^{ins}, for

$$n = 0, \pm 1, \pm 2, \ldots,$$

are mutually orthogonal in L_2:

$$\int_0^{2\pi} e\, e^{-inx} dx = \begin{cases} 0 & m \neq n \\ 2\pi & m = n \end{cases}$$

It follows from this that the functions $e_n\,(n = 0, \pm 1, \pm 2, \ldots)$ defined by $e_n(x) = e^{inx}/\sqrt{2\pi}$ form an orthonormal set in L_2. For any function f in L_2, the numbers

$$c_n = (f, e_n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-inx}\, dx \qquad (6)$$

are its classical Fourier coefficients, and Bessel's inequality takes the form

$$\sum_{n=-\infty}^{\infty} |c_n|^2 \leq \int_0^{2\pi} |f(x)|^2\, dx$$

It is a fact of very great importance in the theory of Fourier series that the orthonormal set $\{e_n\}$ is complete in L_2. As we have seen in Theorem F, the completeness of $\{e_n\}$ is equivalent to the assertion that for every f in L_2, Bessel's inequality can be strengthened to Parseval's equation:

$$\sum_{n=-\infty}^{\infty} |c_n|^2 = \int_0^{2\pi} |f(x)|^2 \, dx$$

Theorem F also tells us that the completeness of $\{e_n\}$ is equivalent to the statement that each f in L_2 has a Fourier expansion:

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad (7)$$

It must be emphasized that this expansion is not to be interpreted as saying that the series converges pointwise to the function. The meaning of (7) is that the partial sums of the series, that is, the vectors f_n in L_2 defined by

$$f_n(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=-n}^{n} c_k e^{ikx} \qquad (7)$$

converge to the vector f in the sense of L_2:

$$\|f_n - f\| \to 0$$

This situation is often expressed by saying that f is the limit in the mean of the f_n's. We add one final remark to our description of this portion of the theory of Fourier series.

If f is an arbitrary function in L_2 with Fourier coefficients c_n, defined by (6), then Bessel's inequality tells us that the series $\sum_{n=-\infty}^{\infty} |c_n|^2$ converges.

The celebrated Riesz-Fischer theorem asserts the converse: if c_n $(n = 0, \pm 1, \pm 2, \dots)$ are given complex numbers for which $\sum_{n=-\infty}^{\infty} |c_n|^2$converges, then there exists a function f in L_2, whose Fourier coefficients are the c_n's.

If we grant the completeness of L_2 as a metric space, this is very easy to prove.

All that is necessary is to use the c_n's to define a sequence of f_n's in accordance with (8). The functions $\frac{e^{inx}}{\sqrt{2\pi}}$ form an orthonormal set, so for $m > n$ we have

$$\|f_m - f_n\|^2 = \sum_{|k|=n+1}^{m} |c_k|^2 \qquad (9)$$

By the convergence of $\sum_{n=-\infty}^{\infty}|c_n|^2$ the sum on the right of (9) can be made as small as we please for all sufficiently large n and all $m > n$.

This tells us that the f_n's form a Cauchy sequence in L_2; and since $L_2$is complete, there exists a function f in L_2 such that $f_n \to f$.

This function f is given by (7), and the c_n's are clearly its Fourier coefficients. It is apparent from these remarks that the essence of the Riesz-Fischer theorem lies in the completeness of L_2 as a metric space.

We shall have use for one further item in the general theory of orthonormal sets, namely, the Gram-Schmidt orthogonalization process.

Suppose that $\{x_1, x_2, \dots, x_n, \dots\}$ is a linearly independent set in a Hilbert space H. The problem is to exhibit a constructive procedure for converting this set into a corresponding orthonormal set $\{e_1, e_2, \dots, e_n, \dots\}$ with the property that for each n the linear subspace of H spanned by $\{e_1, e_2, \dots, e_n\}$ is the same as that spanned by $\{x_1, x_2, \dots, x_n\}$.

Our first step is to normalize z,;—which is necessarily non- zero—by putting

$$e_1 = \frac{x_1}{\|x_1\|}$$

The next step is to subtract from z2 its component in the direction of e_1 to obtain the vector $x_2 - (x_2 e_1)e_1$ orthogonal to e_1, and then to normalize this by putting

$$e_2 = \frac{x_2 - (x_2 e_1)e_1}{\|x_2 - (x_2 e_1)e_1\|}$$

We observe that since x_2 is not a scalar multiple of x_1, the vector $x_2 - (x_2e_1)e_1$ is not zero, so the definition of e_2 is valid. Also, it is clear that e_2 is a linear combination of x_1 and x_2, and that x_2 is a linear combination of e_1 and e_2. The next step is to subtract from x_2 its components in the directions of e_1 and e_2 to obtain a vector orthogonal to e_1 and e_2, and then to normalize this by putting

$$e_3 = \frac{x_3 - (x_3e_1)e_1 - (x_3e_2)e_2}{\|x_3 - (x_3e_1)e_1 - (x_3e_2)e_2\|}$$

If this process is continued in the same way, it clearly produces an orthonormal set $\{e_1, e_2, \dots, e_n, \dots\}$ with the required property.

Example 4. Many orthonormal sets of great interest and importance in analysis can be obtained conveniently by applying the Gram-Schmidt process to sequences of simple functions.

(a) In the space Ly associated with the interval $[-1,1]$, the functions x_n $(n = 0, 1, 2, \dots)$ are linearly independent. If we take these functions to be the x_n's in the Gram-Schmidt process, then the e_n's are the normalized Legendre polynomials.

(b) Consider the space L_2 over the entire real line. If the x_n's here are taken to be the functions $x_n e^{-\frac{x^2}{2}}$ $(n = 0, 1, 2, \dots)$, then the corresponding e_n's are the normalized Hermite functions.

(c) Consider the space L_2. associated with the interval $[0, +\infty)$. If the x_n's are the functions $x_n e^{-x}$ $(n = 0, 1, 2, \dots)$, then the e_n's are the normalized Laguerre functions.

Each of the orthonormal sets described in the above example can be shown to be complete in its corresponding Hilbert space. The analysis involved in a detailed study of these matters is quite complicated and has no proper place in the present book. The reader should recognize, however—and this is our only reason for mentioning the material in Examples 3 and 4—that the theory of Hilbert spaces does have significant contacts with many solid topics in analysis.

Problems

1. Let $\{e_1, e_2, \dots, e_n\}$ bea finite orthonormal set in a Hilbert space H, and let x be a vector in H. If $\alpha_1, \alpha_2, \dots, \alpha_n$ are arbitrary scalars, show that $\|x - \sum_{i=1}^{n} a_i e_i\|$ attains its minimum value $\Leftrightarrow$

$$\alpha_i = (x, e_i)$$

for each i. (Hint: expand $\|x - \sum_{i=1}^{n} a_i e_i\|$, add and subtract $\sum_{i=1}^{n} |(x, e_i)|^2$, and obtain an expression of the form $\sum_{i=1}^{n} |(x, e_i) - \alpha_i|^2$ {? in the result.)
2. Show that the orthonormal sets described in Examples 1 and 2 are complete.
3. Show that every orthonormal set in a Hilbert space is contained in some complete orthonormal set, and use this fact to give an alternative proof of Theorem 53-B.
4. Prove that a Hilbert space H is separable = every orthonormal set in His countable.
5. Show that an orthonormal set in a Hilbert space is linearly independent, and use this to prove that a Hilbert space is finite-dimensional $\Leftrightarrow$ every complete orthonormal set is a basis.
6. Prove that any two complete orthonormal sets in a Hilbert space H have the same cardinal number. This cardinal number is called the orthogonal dimension of H (if H has no complete orthonormal sets, its orthogonal dimension is said to be 0).

UNIT - 3

3.1 THE CONJUGATE SPACE H*

We pointed out in the introduction to this chapter that one of the fundamental properties of a Hilbert space H is the fact that there is a natural correspondence between the vectors in H and the functionals in H*. Our purpose in this section is to develop the features of this correspondence which are relevant to our work with operators in the rest of the chapter.

Let y be a fixed vector in H, and consider the function f_y defined on$H$ by $f_y(x) = (x, y)$. It is easy to see that f_y is linear, for

$$f_y(x_1 + x_2) = (x_1 + x_2, y)$$

$$= (x_1, y) + (x_2, y)$$

$$= f_y(x_1) + f_y(x_2)$$

and

$$f_y(\alpha x) = (\alpha x, y)$$

$$= \alpha(x, y)$$

$$= \alpha f_y(x)$$

Further, f_y is continuous and is therefore a functional, for Schwarz's inequality gives

$$|f_y(x)| = |(x, y)|$$

$$\leq \|x\|\|y\|$$

which shows that$\|f_y\| \leq \|y\|$. Even more, equality is attained here, that is, $\|f_y\| = \|y\|$. This is clear if $y = 0$; and if $y \neq 0$, it follows from

$$\|f_y\| = \sup\{|f_y(x)| : \|x\| = 1\}$$

$$\geq \left| f_y \left(\frac{y}{\|y\|} \right) \right|$$

$$= \left| \left(\frac{y}{\|y\|}, y \right) \right|$$

$$= \|y\|$$

To summarize, we have seen that $y \to f_y$, is a norm-preserving mapping of H into H*. This observation would be of no more than passing interest if it were not for the fact that every functional in H* arises in just this way.

Theorem 3.1 : Let H be a Hilbert space, and let f be an arbitrary functional in H*. Then there exists a unique vector y in H such that

$$f(x) = (x, y) \tag{1}$$

for every $x$in H.

Proof.

It is easy to see that if such a y exists, then it is necessarily unique.

For if we also have $f(x) = (x, y')$ for all x, then $(x, y') = (x, y)$ and $(x, y' — y) = 0$ for all x; and since 0 is the only vector orthogonal to every vector, this implies that $y' — y = 0$or$y' = y$.

We now turn to the problem of showing that y does exist. If $f = 0$, then it clearly suffices to choose $y = 0$.

We may therefore assume that $f \neq 0$. The null space M of f is thus a proper closed linear subspace of H, and by Theorem, there exists a non-zero vector yo which is orthogonal to M.

We show that if a is a suitably chosen scalar, then the vector $y = \alpha y_0$ meets our requirements. We first observe that no matter what a may be, (1) is true for every x in M;

for $f(x) = 0$ for such an x, and since x is orthogonal to y_0, we also have $(x, y) = 0$. This allows us to focus our attention on choosing α in such a way that (1) is true for $x = y_0$. The condition this imposes on α is that

$$f(y_0) = (y_0, \alpha y_0) = \overline{\alpha}\|y_0\|^2$$

We therefore choose a to be $\overline{f(y_0)}/\|y_0\|^2$, and it follows that (1) is true for every x in M and for $x = y_0$.

It is easily seen that each x in H can be written in the form $x = m + \beta y_0$ with m in M: all that is necessary is to choose β in such a way that $f(x - \beta y_0) = f(x) - \beta f(y_0) = 0$, and this is accomplished by putting $\beta = f(x)/f(y_0)$.

Our conclusion that (1) is true for every x in H now follows at once from

$$f(x) = f(m + \beta y_0)$$

$$= f(m) + \beta f(y_0)$$

$$= (m, y) + \beta(y_0, y)$$

$$= (m + \beta y_0, y)$$

$$= (x, y)$$

This result tells us that the norm-preserving mapping of H into H* defined by

$$y \to f_y \quad where \quad f_y(x) = (x, y) \qquad (2)$$

is actually a mapping of H onto H*, It would be pleasant if (2) were also a linear mapping., This is.not quite true, however, for

$$f_{y_1+y_2} = f_{y_1} + f_{y_2} \quad and \quad f_{\alpha y} = \overline{\alpha} f_y \qquad (3)$$

It is an easy consequence of (3) that the mapping (2) is an isometry, for $\|f_x - f_y\| = \|f_{x-y}\| = \|x - y\|$. We state several interesting additional facts about this mapping (and what it enables us to do) in the problems, and we leave their verification to the reader. It should be remembered, however, that the real significance of this entire circle of ideas lies in its influence on the theory of the operators on H. We begin the treatment of these matters in the next section.

Problems

1. Verify relations (3).
2. Let H be a Hilbert space, and show that H* is also a Hilbert space with respect to the inner product defined by $(f_x, f_y) = (y, x)$. In just the same way, the fact that H* is a Hilbert space implies that H** is a Hilbert space whose inner product is given by $(F_f, F_g) = (g, f)$.
3. Let H be a Hilbert space. We have two natural mappings of H into H**, the second of which is onto: the Banach space natural imbedding $f \to F_x$where $F_x(f) = f(x)$, and the product mapping $x \to f_x \to F_{f_x}$, where $f_x(y) = (y, x)$ and$F_{f_x}(f) = (f, f_x)$. Show that these mappings are equal, and conclude that H is reflexive. Show also that $(F_x, F_y) = (x, y)$.

3.2 THE ADJOINT OF AN OPERATOR

Throughout the rest of this chapter, we focus our attention on a fixed but arbitrary Hilbert space H, and unless we specifically state otherwise, it is to be understood that H is the context for all our discussions and theorems.

Let T be an operator on H. We saw in Sec. 51 that T gives rise to an operator T* (its conjugate) on H*, where T* is defined by

$$(T^* f)x = f(Tx)$$

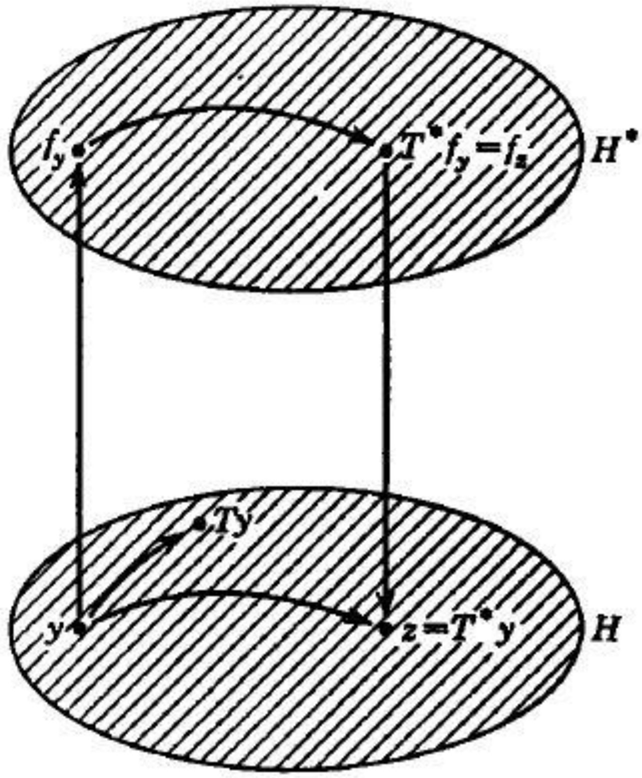

Fig. 3.1**. The conjugate and the adjoint of *T***

We also saw that the mapping T → T* is an isometric isomorphism of $\mathcal{C}$(H) into $\mathcal{C}$(H*) which reverses products and preserves the identity transformation.

In the same way, T* gives rise to an operator T** on H**; and since H is reflexive, it follows that T** = T when H** is identified with H by means of the natural imbedding.

These statements depend only on the fact that H is a reflexive Banach space.

We now bring its Hilbert space character into the picture, and we use the natural correspondence between H and H* discussed in the previous section to pull T* down to H. The details of this procedure are as follows (see Fig. 3.1).

Let y be a vector in H, and f_y its corresponding functional in H*; operate with T* on Fig. 3.1. The conjugate and the f_y to obtain a functional f_x = T*$f_y$and adjoint of T return to its corresponding vector z in H.

There are three mappings under consideration here, and we are forming their product:

$$y \to f_y \to T^* f_y = f_z \to z \quad (1)$$

We write z = T*y, and we call this new mapping T* of H into itself the adjoint of T. The same symbol is used for the adjoint of 7 as for its conjugate because these two mappings are actually the same if H and H* are identified by means of the natural correspondence.

It is easy to keep track of whether T* signifies the conjugate or the adjoint of T by noticing whether it operates on functionals or on vectors.

The action of the adjoint can be linked more closely to the structure of H by observing that for every vector x we have (T*f_y)x = f_x,(Tx) = (Tx,y) and (T*f_y)x = f_y(x) = (x,z) = (x,T*y), so that

$$(Tx, y) = (x, T^* y) \quad (2)$$

Forall$x$and y. Equation (2) ismuch more than merely a property of the adjoint of T, for it uniquely determines this adjoint.

The proof is simple: if T' is any mapping of H into itself such that $(Tx, y) = (x, T'y)$ for all x and y, then $(x, T'y) = (x, T^* y)$ for all x, so $T'y = T^* y^{-1}$ and since the latter is true for all $y, T' = T^*$.

Our remarks in the above paragraph have shown that to each operator T on H there corresponds a unique mapping T* of H into itself (called the adjoint of T) which satisfies relation ‘(2) for all x and y.

There is a more direct but less natural approach to these ideas, one which avoids any reference to the conjugate of T.

If y is fixed, it is clear that the expression (Tx, y) is a scalar-valued continuous linear function of x. By Theorem, there exists a unique vector z such that $(Tx, y) = (x, z)$ for all x. We now write z = T*y, and since y is arbitrary, we again have relation (2) for all x and$y$. The fact that T* is uniquely determined by (2) follows just as before.

The principal value of our approach to the definition of the adjoint (as opposed to that just mentioned) lies in the motivation it provides for considering adjoints at all.

We can express this by emphasizing that an operator on a Banach space always has a conjugate which operates on the conjugate space; and when the Banach space happens to be a Hilbert space, then, as we have seen, the natural correspondence discussed in the previous section makes it almost inevitable that we regard the conjugate as an operator on the space itself. Once the definition of the adjoint is fully understood, however, there is no further need to mention conjugates. All our future work with adjoints will be based on Eq. (2), and from this point on, the symbol T* will always signify the adjoint of T (and never its conjugate).

As our first step in exploring the properties of adjoints, we verify that T* actually is an operator on H (all we know so far is that it maps A into itself). For any y and z, and for all x, we have

$$\big(x, T^*(y+z)\big) = (Tx, y+z)$$

$$= (Tx, y) + (Tx, z)$$

$$= (x, T^*y) + (x, T^*z)$$

$$= (x, T^*y + T^*z)$$

so

$$T^*(y+z) = T^*y + T^*z$$

The relation $T^*(\alpha y) = \alpha T^* y$

is proved similarly, so T* is linear. It remains to be seen that T* is continuous; and to prove this, we note that

$$\|T^*y\|^2 = (T^*y, T^*y)$$

$$= (TT^*y, y)$$

$$\leq \|TT^*y\|\|y\|$$

$$\leq \|T\|\|T^*y\|\|y\|$$

implies that $\|T^*y\| \leq \|T\|\|y\|$ for all y, so

$$\|T^*\| \leq \|T\|$$

These facts tell us that T → T* is a mapping of $\mathcal{C}$(H) into itself. This mapping is called the adjoint operation on $\mathcal{C}$(H).

Theorem 3.2. The adjoint operation T→T* on $\mathcal{C}$(H) has the following properties:

(1) $(T_1 + T_2) = T_1^* + T_2^*$

(2) $(\alpha$T)* = αT*;

(3) $(T_1T_2)^* = T_1^*T_2^*$

(4) T**=T

(5) $\|T^*\| = \|T\|$

(6) $\|T^*T\| = \|T\|^2$

Proof.

The arguments used in proving (1) to (4) are all essentially the same. As an illustration of the method, we observe that (3) follows from the fact that for all x and y we have

$$(x, (T_1T_2)^*y) = (T_1T_2x, y)$$

$$= (T_2x, T_1^*y)$$

$$= (x, T_2^*T_1^*y)$$

To prove (5), we note that we already have$\|T^*\| \leq \|T\|$; and if we apply this to T* instead of T and use (4), we obtain $\|T\| = \|T^{**}\| \leq \|T^*\|$. Half of (6) follows from (5) and the inequality (5), for

$$\|T^*T\| \leq \|T^*\|\|T\|$$

$$= \|T\|\|T\|$$

$$= \|T\|^2$$

and the fact that $\|T\|^2 \leq \|T^*T\|$is an immediate consequence of

$$\|Tx\|^2 = (Tx, Tx)$$

$$= (T^*Tx, x)$$

$$= \|T^*Tx\|\|x\|$$

$$\leq \|T^*T\|\|x\|^2$$

The presence of the adjoint operation is what distinguishes the theory of the operators on H from the more general theory of the operators on a reflexive Banach space. In the next three sections, we use this operation as a tool by means of which we single out for special study certain types of operators on H whose theory is particularly complete and satisfying.

Problems

1. Prove parts (1), (2), and (4) of Theorem A.
2. Show that the adjoint operation is one-to-one onto as a mapping of $\mathcal{C}$(H) into itself.
3. Show that 0* = 0 and I* = I. Use the latter to show that if T is non-singular, then T* is also non-singular, and that in this case $(T^*)^{-1} = (T^{-1})^*$.
1. 4, Show that $\|T^*T\| = \|T\|^2$.

3.3 SELF-ADJOINT OPERATORS

There is an interesting analogy between the set $\mathcal{C}$(H) of all operators on our Hilbert space H and the set C of all complex numbers. This can be summarized by observing that each is a complex algebra together with a mapping of the algebra onto itself (T→T* and $z \to \bar{z}$) and that these mappings have similar properties. We shall see that this analogy is quite useful as an intuitive guide to the study of the operators on H. The most significant difference between these systems is that multi- plication in the algebra $\mathcal{C}$(H) is in general non-commutative, and it will become clear as we proceed that this is the primary source of the much greater structural complexity of $\mathcal{C}$(H).

The most important subsystem of the complex plane is the real line, which is characterized by the relation $z = \overline{z}$. By analogy, we consider those operators A on H which equal their adjoints, that is, which satisfy the condition A = A*. Such an operator is said to be self-adjoint. The self-adjoint operators on H are evidently those which are related in the simplest possible way to their adjoints.

We know that 0* = 0 and I* = I, so 0 and I are self-adjoint. It A_1 and A_2 are self-adjoint, and if α and β are real numbers, then

$$(\alpha A_1 + \beta A_2)^* = \overline{\alpha} A_1^* + \overline{\beta} A_2^*$$

$$= \alpha A_1 + \beta A_2$$

shows that $\alpha A_1 + \beta A_2$ is also self-adjoint. Further, if $\{A_n\}$ isa sequence of self-adjoint operators which converges to an operator A, then it is easy to see that A is also self-adjoint; for

$$\|A - A^*\| \leq \|A - A_n\| + \|A_n - A_n^*\| + \|A_n^* - A^*\|$$

$$= \|A - A_n\| + \|(A_n - A)^*\|$$

$$= \|A - A_n\| + \|A_n - A\|$$

$$= 2\|A_n - A\| \to 0$$

shows that A − A* = 0, so A = A*. These remarks yield our first theorem.

Theorem 3.3 : The self-adjoint operators in $\mathcal{C}$(H) form a closed real linear subspace of $\mathcal{C}$(H)—and therefore a real Banach space which contains the identity transformation.

The reader will notice that we have said nothing here about the product of two self-adjoint operators. Very little is known about such products, and the following simple result represents almost the extent of our information.

Theorem 3.4 : If A_1 and $A_2$are self-adjoint operators on H, then their product $A_1A_2$is self-adjoint $\Leftrightarrow A_1A_2 = A_2A_1$.

Proof.

This is an obvious consequence of

$$(A_1A_2)^* = A_2^*A_1^* = A_2A_1$$

The order properties of self-adjoint operators are more interesting, and we devote the remainder of the section to establishing some of the simpler facts in this direction.

If T is an arbitrary operator on H, it is easy to see that

$$T = 0 \Leftrightarrow (Tx, y) =$$

for all$x$ and y. It is also clear that T = 0 = (Tx,x) = 0 for all$x$. We shall need the converse of this implication.

Theorem 3.5 If T is an operator on H for which (Tx,x) = 0 for all x, then T = 0.

Proof.

It suffices to show that (Tx,y) = 0 for any x and y, and the proof of this depends on the following easily verified identity:

$$(T(\alpha x + \beta y), \alpha x + \beta y) = |\alpha|^2(Tx, x) - |\beta|^2(Ty, y) = \alpha\overline{\beta}(Tx, y) + \overline{\alpha}\beta(Ty, x) \quad (1)$$

We first observe that by our hypothesis, the left side of (1)—and therefore the right side as well—equals 0 for all a and g. If we put α = 1 andβ = 1, then (1) becomes

$$(Tx, y) + (Ty, x) = 0 \quad (2)$$

and if we put $\alpha = i$ and $\beta = 1$, we get

$$i(Tx, y) + i(Ty, x) = 0 \quad (3)$$

Dividing (3) by < and adding the result to (2) yields $2(Tx,y)=0$ so $(Tx,y) = 0$ and the proof is complete.

It is worth emphasizing that this proof makes essential use of the fact that the scalars are the complex numbers (and not merely the real numbers).

We now apply this result to proving our next theorem, which indicates that self-adjoint operators are linked to real numbers by stronger ties than might be suspected from the loose analogy that led to their definition.

Theorem 3.6 : An operator T on H is self-adjoint $\Leftrightarrow$ (Tx,x) is real for all x.

Proof :

If T is self-adjoint, then

$$\overline{(Tx,x)} = (x,Tx)$$

$$= (x,T^*x)$$

$$= (Tx,x)$$

shows that (Tx,x) is real for all x. On the other hand, if (Tx,x) is real for all x, then $(Tx,x) = \overline{(Tx,x)} = \overline{(x,T^*x)} = (T^*x,x)$ or

$$([T-T^*]x,x) = 0$$

for all x. By Theorem C, this implies that T−T* = 0, so T = T*.

This theorem enables us to define a respectable and useful order relation on the set of all self-adjoint operators. If A_1 and A_2 are self- adjoint, we write $A_1 \leq A_2$ if $(A_1x,x) \leq (A_2x,x)$ for all x. The main elementary facts about this relation are summarized in

Theorem 3.7 : The real Banach space of all self-adjoint operators on H is a partially ordered set whose linear structure and order structure are related by the following properties:

(1) if $A_1 \leq A_2$, then $A_1 + A \leq A_2 + A$ for every A;

(2) if $A_1 \leq A_2$ and $\alpha > 0$, then $\alpha A_1 \leq \alpha A_2$.

Proof.

The relation in question is obviously reflexive and transitive (see Sec. 8). To show that it is also antisymmetric, we assume that $A_1 \leq A_2$ and $A_2 \leq A_1$.

This implies at once that $([A_1 - A_2]x, x) = 0$ for all x, so by Theorem C, $A_1 - A_2 = 0$ and $A_1 = A_2$.

The proofs of properties (1) and (2) are easy. For instance, if $A_1 \leq A_2$, so that $(A_1 x, x) \leq (A_2 x, x)$ for all x, then $(A_1 x, x) + (Ax, x) \leq (A_2 x, x) + (Ax, x)$ or $([A_1 - A]x, x)$ for all x, so $A_1 + A \leq A_2 + A$. The proof of (2) is similar.

A self-adjoint operator A is said to be positive if $A \geq 0$, that is, if $(Ax, x) \geq 0$ for all x. It is clear that 0 and I are positive, as are T*T and TT* for an arbitrary operator T.

Theorem 3.8 : If A is a positive operator on H, then $I + A$ is non-singular. In particular, I + T*T and I + TT* are non-singular for an arbitrary operator T on H.

Proof.

We must show that I + A is one-to-one onto as a mapping of H into itself. First, it is one-to-one, for

$$(I + A)x = 0 \Rightarrow Ax = -x \Rightarrow (Ax, x) = (-x, x) = -\|x\|^2 \geq 0 \Rightarrow x = 0.$$

We next show that the range M of $I + A$ is closed. It follows from $\|(I + A)x\|^2 = \|x\|^2 + \|Ax\|^2 + 2(Ax, x)$and the assumption that A is positive—that $\|x\| \leq \|(I + A)x\|$. By this inequality and the com- pleteness of H, M is complete and therefore closed.

We conclude the proof by observing that $M = H$; for otherwise there would exist a non-zero vector x_0 orthogonal to M, and this would contradict the fact that$(x_0, [I + A]x_0) = 0 \Rightarrow \|x_0\|^2 = -(Ax_0, x_0) \leq 0 \Rightarrow x_0 = 0$.

If the reader wonders why we fail to show that the partially ordered set of all self-adjoint operators is a lattice, the reason is simple: it isn't true. As a matter of fact, this system is about as far from being a lattice as a partially ordered set can be, for it can be shown that two operators in the set have a greatest lower bound = they are comparable. This whole situation is intimately related to questions of commutativity for algebras of operators and is too complicated for us to explore here. For further details, see Kadison [22].

Problems

1. Define a new operation of 'multiplication' for self-adjoint operators by$A_1 \circ A_2 = (A_1A_2 + A_2A_1)/2$, and note that $A_1 \circ A_2$ is always self-adjoint and that it equals A_1A_2whenever A_1 and A_2 commute. Show that this operation has the following properties:

$$A_1 \circ A_2 = A_2 \circ A_2$$

$$A_1 \circ (A_2 + A_3) = A_1 \circ A_2 + A_1 \circ A_3$$

$$\alpha(A_1 \circ A_2) = (\alpha A_1) \circ A_2 = A_1 \circ (\alpha A_2)$$

and $A \circ I = I \circ A = A$. Showalso that $A_1 \circ (A_2 \circ A_3) = (A_1 \circ A_2) \circ A_3$ whenever A_1 and A_2commute.

2. If T is any operator on H, it is clear that$|(Tx, x)| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2$; so if $H \neq \{0\}$, we have $\sup\{|Tx, x|/\|x\|^2 : x \neq 0\} \leq \|T\|$. Prove that if T is self-adjoint, then equality holds here. (Hint: write$a = \sup\{|Tx, x|/\|x\|^2 : x \neq 0\} = \sup\{|Tx, x|/\|x\| = 1\}$, and show that$\|Tx\| \leq a$ whenever $\|x\| = 1$ by putting $b = \|Tx\|^{1/2}$if$Tx \neq 0$—and considering

$$4\|Tx\|^2 = \big(T(bx + b^{-1}Tx), (bx + b^{-1}Tx)\big) - \big(T(bx + b^{-1}Tx), (bx + b^{-1}Tx)\big)$$

$$\leq a[\|bx + b^{-1}Tx\|]^2 + \|bx + b^{-1}Tx\|$$

$$= 4a\|Tx\|$$

3.4 NORMAL AND UNITARY OPERATORS

An operator N on H is said to be normal if it commutes with its adjoint, that is, if NN* = N*N. The reason for the importance of normal operators will not become clear until the next chapter. We shall see that they are the most general operators on H for which a simple and revealing structure theory is possible. Our purpose in this section is to present a few of their more elementary properties which are necessary for our later work.

It is obvious that every self-adjoint operator is normal, and that if N is normal and a is any scalar, then aN is also normal. Further, the limit N of any convergent sequence {N;.} of normal operators is normal; for we know that $N_k{}^* \to N^*$, so

$$\|NN^* - N^*N\| \leq \|NN^* - N_kN_k^*\| + \|N_kN_k^* - N_k^*N_k\| + \|N_k^*N_k - N^*N\|$$

$$= \|NN^* - N_kN_k^*\| + \|N_k^*N_k - N^*N\| \to 0$$

which implies that NN* − N*N = 0. These remarks prove

Theorem 3.9 : The set of all normal operators on H is a closed subset of $\mathcal{C}$(H) which contains the set of all self-adjoint operators and is closed under scalar multiplication.

It is natural to wonder whether the sum and product of two normal operators are necessarily normal. They are not, but nevertheless, we can say a little in this direction.

Theorem 3.10 : If N_1 and N_2 are normal operators on H with the property that either commutes with the adjoint of the other, then $N_1 + N_2$ and N_1N_2 are normal.

Proof.

It is clear by taking adjoints that

$$N_1N_2^* = N_2^*N_1 \Leftrightarrow N_2N_1^* = N_1^*N_2$$

so the assumption implies that each commutes with the adjoint of the other. To show that $N_1 + N_2$ is normal under the stated conditions, we have only to compare the results of the following computations:

$$(N_1 + N_2)(N_1 + N_2)^* = (N_1 + N_2)(N_1^* + N_2^*)$$

$$= N_1 N_1^* + N_1 N_2^* + N_2^* N_1 + N_2 N_2^*$$

$$and\ (N_1 + N_2)^*(N_1 + N_2) = (N_1^* + N_2^*)(N_1 + N_2)$$

$$= N_1^* N_1 + N_1^* N_2 + N_2^* N_1 + N_2^* N_2$$

The fact that $N_1 N_2$ is normal follows similarly from

$$N_1 N_2 (N_1 N_2)^* = N_1 N_2 N_2^* N_1^*$$

$$= N_1 N_2^* N_2 N_1^*$$

$$= N_2^* N_1 N_1^* N_2$$

$$= N_2^* N_1^* N_1 N_2$$

$$= (N_1 N_2)^* N_1 N_2$$

By definition, a self-adjoint operator A is one which satisfies the identity A*x = Ax. Many properties of self-adjoint operators do not depend on this, but only on the weaker identity $\|A^* x\| = \|Ax\|$. Our next theorem shows that all such properties are shared by normal operators.

Theorem 3.11 : An operator T on H is normal $\|T^* x\| = \|Tx\|$ for every x.

Proof.

We have,

$$\|T^* x\| = \|Tx\| \Leftrightarrow \|T^* x\|^2 = \|Tx\|^2$$

$$\Leftrightarrow (T^* x, T^* x) = (Tx, Tx)$$

$$\Leftrightarrow (TT^* x, x) = (T^* T x, x)$$

$$\Leftrightarrow ([TT^* - T^* T]x, x) = 0$$

The following consequence of this result will be useful in our later work.

Theorem 3.12 : If N is a normal operator on H, then$\|N^2\| = \|N\|^2$.

Proof.

The preceding theorem shows that

$$\|N^2x\| = \|NNx\| = \|N^*Nx\|$$

for every x, and this implies that $\|N^2x\| = \|N^*N\|$. By Theorem 56-A, we have $\|N^*N\| = \|N\|^2$, so the proof is complete.

We know that any complex number z can be expressed uniquely in the form $z = a + ib$ where a and b are real numbers, and that these real numbers are called the real and imaginary parts of z and are given by $a = (z + \overline{z})/2$ and$b = (z - \overline{z})/2i$.

The analogy between general operators and complex numbers, and between self-adjoint operators and real numbers, suggests that for an arbitrary operator T on H we form $A_1 = (T + T^*)/2$ and $A_2 = (T - T^*)/2i$. A_1 and A_2 are clearly self- adjoint, and they have the property that $T = A_1 + iA_2$.

The unique- ness of this expression for T follows at once from the fact that

$$T^* = A_1 - iA_2$$

The self-adjoint operators A_1 and A_2 are called the real part and the tmaginary part of T.

We emphasized earlier that the complicated structure of $\mathcal{C}(H)$ is due in large part to the fact that operator multiplication is in general non-commutative.

Since our future work will be focused mainly on normal operators, it is of interest to see as the following theorem shows that the existence of non-normal operators can be traced directly to the non-commutativity of self-adjoint operators.

Theorem 3.13 : If T is an operator on H, then T is normal $\Leftrightarrow$ its real and imaginary parts commute.

Proof.

If A_1 and A_2 are the real and imaginary parts of T, so that $T = A_1 + iA_2$ and $T^* = A_1 - iA_2$, then

$$TT^* = (A_1 + iA_2)(A_1 - iA_2)$$

$$= A_1^2 + A_2^2 + i(A_2A_1 - A_1A_2)$$

and

$$T^*T = (A_1 - iA_2)(A_1 + iA_2)$$

$$= A_1^2 + A_2^2 + i(A_1A_2 - A_2A_1)$$

It is clear that if $A_1A_2 = A_2A_1$, then $TT^* = T^*T$. Conversely, if $TT^* = T^*T$, then $A_1A_2 - A_2A_1 = A_2A_1 - A_1A_2$, so $2A_1A_2 = 2A_2A_1$ and $A_1A_2 = A_2A_1$.

Perhaps the most important subsystem of the complex plane after the real line is the unit circle, which is characterized by either of the equivalent identities $|z| = 1$ or $z\bar{z} = \bar{z}z = 1$. An operator U on H which satisfies the equation UU* = U*U =I is said to be unitary.

Unitary operators which are obviously normal are thus the natural analogues of complex numbers of absolute value 1. It is clear from the definition that the unitary operators on H are precisely the non-singular operators whose inverses equal their adjoints. The geometric significance of these operators is best. understood in the light of our next theorem.

Theorem 3.14 : If T is an operator on H, then the following conditions are all equivalent to one another:

(1) $T^*T = T$;

(2) $(Tx, Ty) = (x, y)$ for all x and y;

(3) $\|Tx\| = \|x\|$ for all x.

Proof.

If (1) is true, then $(T^*Tx, y) = (x, y)$or $(Tx, Ty) = (x, y)$for all x and y, so (2) is true; and if (2) is true, then by taking $y = x$ we obtain $(Tx, Tx) = (x, x)$ or $\|Tx\|^2 = \|x\|^2$ for all x, so (3) is true.

The fact that (3) implies (1) is a consequence of Theorem and the follow- ing chain of implications:

$$\|Tx\| = \|x\| \Rightarrow \|Tx\|^2 = \|x\|^2$$

$$\Rightarrow (Tx, Tx) = (x, x)$$

$$\Rightarrow (T^*Tx, y) = (x, y)$$

$$\Rightarrow ([T^*T - I]x, x) = 0$$

An operator on H with property (3) of this theorem is simply an isometric isomorphism of H into itself. That an operator of this kind need not be unitary is easily seen by considering the operator on l_2 defined by

$$T\{x_1, x_2, \dots\} = \{0, x_1, x_2, \dots\}$$

which preserves norms but has no inverse. These ideas lead at once to

Theorem 3.15 : An operator Ton H is unitary ⇔ it is an isometric isomorphism of H onto itself.

Proof.

If T is unitary, then we know from the definition that it is onto; and since by Theorem F it preserves norms, it is an isometric isomorphism of H onto itself.

Conversely, if T is an isometric isomorphism of H onto itself, then T^{-1} exists, and by Theorem F we have T*T =I. It now follows that (T*T)T^{-1} = IT^{-1}, so T* =T^{-1}) and TT* = T*T = I, which shows that T is unitary.

This theorem makes quite clear the nature of unitary operators: they are precisely those one-to-one mappings of H onto itself which preserve all structure—the linear operations, the norm, and the inner product.

Problems

1. If T is an arbitrary operator on H, and if α and β are scalars such that $|\alpha| = |\beta|$, show that $\alpha T + \beta T^{*}$is normal.
2. If H is finite-dimensional, show that every isometric isomorphism of H into itself is unitary.
3. Show that an operator T onH is unitary ⇔T($\{e_i\}$) is a complete orthonormal set whenever $\{e_i\}$ is.
4. Show that the unitary operators on H form a group.

3.5 PROJECTIONS

According to the definition given in Sec. 50, a projection on a Banach space B is an idempotent operator on B, that is, an operator P with the property that $P^2 = P$. It was proved in that section that each projection P determines a pair of closed lincar subspaces M and N the range and null space of P such that $B = M \oplus N$, and also, conversely, that each such pair of closed linear subspaces M and N determines a projection P with range M and null space N. In this way, there is established a one-to-one correspondence between projections on B and pairs of closed linear subspaces of B which span the whole space and have only the zero vector in common.

The context of our present work, however, is the Hilbert space H, and not a general Banach space, and the structure which H enjoys in addition to being a Banach space enables us to single out for special attention those projections whose range and null space are orthogonal. Our first theorem gives a convenient characterization of these projections.

Theorem 3.16 If P is a projection on H with range M and null space N, then M ⊥ N ⇔ P is self-adjoint; and in this case, $N = M^{\perp}$.

Proof.

Each vector z in H can be written uniquely in the form $z = x + y$ with $x$and $y$in $M$and N. If $M \perp N$, so that $x \perp y$, then P* = P will follow by Theorem 57-C from $(P^{*}z, z) = (Pz, z)$; and this is a consequence of

$$(P^{*}z, z) = (z, Pz) = (z, x) = (x + y, x) = (x, x) + (y, x) = (x, x)$$

and $(Pz, z) = (x, z) = (x, x + y) = (x, x) + (x, y) = (x, x)$. If, conversely, P* = P, then the conclusion that $M \perp N$ follows from the fact that for any x and y in M and N we have

$$(x, y) = (Px, y) = (x, P^{*}y) = (x, Py) = (x, 0) = 0$$

All that remains is to see that if $M \perp N$, then$N = M^{\perp}$. It is clear that $N \subseteq M^{\perp}$ and if N is a proper subset of $M^{\perp}$, and therefore a proper closed linear subspace of the Hilbert space $M^{\perp}$, then Theorem implies that there exists a non-zero vector z_0 in $M^{\perp}$ such that $z_0 \perp N$. Since $z_0 \perp M$and $z_0 \perp N$and since $H = M \oplus N$, it follows that $z_0 \perp H$. This is impossible, so we conclude that $N = M^{\perp}$.

A projection on H whose range and null space are orthogonal is sometimes called a perpendicular projection.

The only projections considered in the theory of Hilbert spaces are those which are perpendicular, so it is customary to omit the adjective and to refer to them simply as projections. In the light of this agreement and Theorem A, a projection on H can be defined as an operator P which satisfies the conditions $P^2 = P$ and P* = P. The operators 0 and I are projections, and they are distinct ⇔ $H \neq \{0\}$.

The great importance of the projections on H rests mainly on Theorem which allows us to set up a natural one-to-one correspondence between projections and closed linear subspaces. To each projection P there corresponds its range $M = \{Px: x \in H\}$, which is a closed linear subspace; and conversely, to each closed linear subspace M there corresponds the projection P with range M defined by $P(x + y) = x$, where x and y are in M and $M^{\perp}$. Either way, we speak of P as the projection on M.

It is clear that P is the projection on $M \Leftrightarrow I - P$ is the projection on $M^{\perp}$. Also, if P is the projection on M, then

$$x \in M \Leftrightarrow Px = x \Leftrightarrow \|Px\| = \|x\|$$

The first equivalence here was proved in Problem 44-11; and since for every x in H we have

$$\|x\|^2 = \|Px + (I - P)x\|^2 = \|Px\|^2 + \|(I - P)x\|^2 \qquad (1)$$

the non-trivial part of the second is given by the following chain of implications:

$$\|Px\| = \|x\| \Rightarrow \|Px\|^2 = \|x\|^2 \Rightarrow \|(I - P)x\|^2 = 0 \Rightarrow Px = x$$

Relation (1) also shows that $\|Px\| \leq \|x\|$ for every x, so $\|P\| < 1$. Ifz is an arbitrary vector in H, it is easy to see that

$$(Px, x) = (PPx, x) = (Px, P^{*}x) = (Px, Px) = \|Px\|^2 \geq 0 \qquad (2)$$

so P is a positive operator $(0 \leq P)$ in the sense of Sec. 57. Since $I - P$ is also a projection, we also have$0 \leq I - P$ or $P \leq I$, so $0 \leq P \leq I$.

Let T be an operator on H. A closed linear subspace M of H is said to be invariant under T if $T(M) \subseteq M$. When this happens, the restriction of T to M can be regarded as an operator on M alone, and the action of T on vectors outside of M can be ignored. If both M and $M^{\perp}$ are invariant under T, we say that M reduces T, or that T is reduced by M. This situation is much more interesting, for it allows us to replace the study of T as a whole by the study of its restrictions to M and $M^{\perp}$, and it invites the hope that these restrictions will turn out to be operators of some particularly simple type. In the following four theorems, we translate these concepts into relations between M and the projection on M.

Theorem 3.17 : A closed linear subspace M of H is invariant under an operator T$\Leftrightarrow M^{\perp}$ is invariant under T*.

Proof.

Since $M^{\perp\perp} = M$ and T** = T, it suffices by symmetry to prove that if M is invariant under T, then $M^{\perp}$ is invariant under T*.

If y is a vector in $M^{\perp}$, our conclusion will follow from $(x, T^{*}y) = 0$ for all x in M. But this is an easy consequence of $(x, T^{*}y) = (Tx, y)$, for the invariance of M under T implies that (Tx, y) = 0.

Theorem 3.18 : A closed linear subspace M of H reduces an operator T⇔ M is invariant under beth T and T*.

Proof.

This is obvious from the definitions and the preceding theorem.

Theorem 3.19 : If P is the projection on a closed linear subspace M of H, then M is invariant under an operator T⇔TP ⇔PT.

Proof.

If M is invariant under T and x is an arbitrary vector in H, then TPx is in M, so $PTPx = TPx$ and $PTP = TP$.

Conversely, if $TP = PTP$ and x is a vector in M, then $Tx = TPx = PTPx$ is also in M, so M is invariant under T.

Theorem 3.20 : If P is the projection on a closed linear subspace M of H, then M reduces an operator T⇔TP = PT.

Proof.

M reduces T ⇔ M is invariant under T and T*⇔TP = PTP and T*P = PT*P ⇔ TP = PTP and PT = PTP.

The last statement in this chain clearly implies that TP = PT; it also follows from it, as we see by multiplying TP = PT on the right and left by P.

Our next theorem shows how projections can be used to express the statement that two closed linear subspaces of H are orthogonal.

Theorem 3.21 : If P and Q are the projections on closed linear subspaces M and N of H, then$M \perp N \Leftrightarrow PQ = 0 \Leftrightarrow QP = 0$.

Proof.

We first remark that the equivalence of $PQ = 0$ and $QP = 0$ is clear by taking adjoints. If $M \perp N$, so that$N \subseteq M^{\perp}$, then the fact that Qx is in N for every x implies that $PQx = 0$, so $PQ = 0$.

If, conversely, $PQ = 0$, then for every x in N we have $Px = PQx = 0$, so $N \subseteq M^{\perp}$and $M \perp N$.

Motivated by this result, we say that two projections P and Q are orthogonal if $PQ = 0$.

Our final theorem describes the circumstances under which a sum of projections is also a projection.

Theorem 3.22 : If $P_1, P_2, \dots, P_n$ are the projections on closed linear sub spaces $M_1. M_2, \dots, M_n$of H, then $P = P_1 + P_2 + \cdots + P_n$isa projection $\Leftrightarrow$ the P_i's arepatrwise orthogonal (in the sense that$P_i P_j = 0$ whenever $i \neq j$); and in this case, P is the projection on

$$M = M_1 + M_2 + \cdots + M_n$$

Proof.

Since P is clearly self-adjoint, it is a projection = it is idem- potent. If the P,'s are pairwise orthogonal, then a simple computation shows at once that P is idempotent.

To prove the converse, we assume that P is idempotent. Let x be a vector in the range of P_i, so that $x = Px$. Then

$$\|x\|^2 = \|Px\|^2 \leq \sum_{j=1}^{n} \|P_i x\|^2 = \sum_{j=1}^{n} \left(P_j x, x\right) = (P_x, x) = \|Px\|^2 \leq \|x\|^2$$

We conclude that equality must hold all along the line here, so

$$\sum_{j=1}^{n} \|P_i x\|^2 = \|Px\|^2$$

and $\|P_j x\| = 0$ for $j \neq i$.

Thus the range of P_i is contained in the null space of P_j that is $M_i \subseteq M_j^{\perp}$ for every $j \neq i$. This means that $M_i \perp M_j$ whenever $i \neq j$, and our conclusion that the P_i's are pairwise orthogonal now follows from the preceding theorem.

We prove the final statement in two steps. First, we observe that since $\|Px\| = \|x\|$ for every x in M_i each M_i is contained in the range of P, and therefore M is also contained in the range of P. Second, if x is a vector in the range of P, then

$$x = P_x = P_1 x + P_2 x + \cdots + P_n x$$

is evidently in M.

There are many other ways in which the algebraic structure of the set of all projections on H can be related to the geometry of its closed linear subspaces. and several of these are given in the problems below.

The significance of projections in the general theory of operators on H is the theme of the next chapter.

As we shall see, the essence of the matter (the spectral theorem) is that every normal operator is made of projections in a way which clearly reveals the geometric nature of its action on the vectors in H.

Problems

1. If P and Q are the projections on closed linear subspaces M and N of H, prove that PQ is a projection ⇔ PQ = QP. In this case, show that PQ is the projection on M∩ N.
2. If P and Q are the projections on closed linear subspaces M and N of H, prove that the following statements are all equivalent to one another:

   (a) $P \leq Q$;

   (b) $\|Px\| \leq \|Qx\|$for every x;

   (c) $M \subseteq N$;

   (d) $PQ = P$;

   (e) $QP = P$.

   (Hint: the equivalence of (a) and (b) is easy to prove, as is that of (c), (d), and (e); prove that (d) implies (a) by using

   $$(Px, x) = \|Px\|^2 = \|PQx\|^2 \leq \|Qx\|^2 = (Qx, x)$$

   and prove that (b) implies (c) by observing that if x is in M, then $\|x\| = \|Px\| \leq \|Qx\| \leq \|x\|$
3. Show that the projections on H form a complete lattice with respect to their natural ordering as self-adjoint operators. (Compare this situation with that described in the last paragraph of Sec. 57.)

CHAPTER - 4

4.1 DETERMINANTS AND THE SPECTRUM OF AN OPERATOR

Determinants are often advertised to students of elementary mathematics as a computational device of great value and efficiency for solving numerical problems involving systems of linear equations. This is somewhat misleading, for their value in problems of this kind is very limited. On the other hand, they do have definite importance as a theoretical tool. Briefly, they provide a numerical means of distinguishing between singular and non-singular matrices (and operators).

This is not the place for developing the theory of determinants in any detail. Instead, we assume that the reader already knows something about them, and we confine ourselves to listing a few of their simpler properties which are relevant to our present interests.

Let $[\alpha_{ij}]$ be an$n \times n$ matrix. The determinant of this matrix, which we denote by $det([\alpha_{ij}])$, is a scalar associated with it in such a way that

(1) $det([\delta_{ij}]) = 1;$

(2) $det([\alpha_{ij}][\beta_{ij}]) = det([\alpha_{ij}])\, det([\beta_{ij}])$;

(3) $det([\alpha_{ij}]) \neq 0 \Leftrightarrow [\alpha_{ij}]$ is non-singular; and

(4) $det([\alpha_{ij}] - \lambda[\delta_{ij}])$ is a polynomial, with complex coefficients, of degree n in the variable λ.

The determinant function det is thus a scalar-valued function of matrices which has certain properties. In elementary work, the determinant of a matrix is usually written out with vertical bars, as follows,

$$\det([\alpha_{ij}]) = \begin{vmatrix} \alpha_{11} & \alpha_{12} \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} \cdots & \alpha_{2n} \\ \cdots & \cdots\cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} \cdots & \alpha_{nn} \end{vmatrix}$$

and is evaluated by complicated procedures which are of no concern to us here.

We now consider an operator Ton H. If Band B' are bases for H, then the matrices $[\alpha_{ij}]$ and $[\beta_{ij}]$ of T relative to B and B' may be entirely different, but nevertheless they have the same determinant. For we know from the previous section that there exists a non-singular matrix $[\gamma_{ij}]$ such that

$$[\beta_{ij}] = [\gamma_{ij}]^{-1}[\alpha_{ij}][\gamma_{ij}]$$

and therefore, by properties (1), (2), and (3), we have

$$\det([\alpha_{ij}]) = \det\left([\gamma_{ij}]^{-1}[\alpha_{ij}][\gamma_{ij}]\right)$$

$$= \det\left([\gamma_{ij}]^{-1}\right)\det([\alpha_{ij}])\det([\gamma_{ij}])$$

$$= \det\left([\gamma_{ij}]^{-1}\right)\det([\gamma_{ij}])\det([\alpha_{ij}])$$

$$= \det\left([\gamma_{ij}]^{-1}[\gamma_{ij}]\right)\det([\alpha_{ij}])$$

$$= \det([\delta_{ij}])\det([\alpha_{ij}])$$

$$= \det([\alpha_{ij}])$$

This result allows us to speak of the determinant of the operator T, meaning, of course, the determinant of its matrix relative to any basis; and from this point on, we shall regard the determinant function primarily as a scalar-valued function of the operators on H. We at once obtain the following four properties for this function, which are simply translations of those stated above:

(1) $det(I) = I$;

(2) $\det(T_1T_2) = \det(T_1)\det(T_2)$;

(3) $\det(T) \neq 0 \Leftrightarrow T$ is non-singular; and

(4) $det(T - \lambda I)$ is a polynomial, with complex coefficients, of degree n in the variable λ.

We are now in a position to take up once again, and to settle, the problem of the existence of eigenvalues.

Let T be an operator on H. If we recall Problem 44-6, it is clear that a scalar λ is an eigenvalue of T $\Leftrightarrow$ there exists a non-zero vector x such that $(T-\lambda I)x = 0 \Leftrightarrow T-\lambda I$ is singular $\Leftrightarrow \det(T-\lambda I) = 0$. The eigenvalues of T are therefore precisely the distinct roots of the equation

$$\det(T-\lambda I) = 0 \qquad (1)$$

which is called the characteristic equation of T. It may illuminate matters somewhat if we choose a basis B for H, find the matrix $[\alpha_{ij}]$ of T relative to B, and write the characteristic equation in the extended form

$$\begin{vmatrix} \alpha_{11}-\lambda & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22}-\lambda & \cdots & \alpha_{2n} \\ \cdots & & \cdots\cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn}-\lambda \end{vmatrix} = 0$$

Our search for eigenvalues of T is reduced in this way to a search for roots of Eq. (1). Property (4') tells us that this is a polynomial equation, with complex coefficients, of degree n in the complex variable λ. We now appeal to the fundamental theorem of algebra, which guarantees that an equation of this kind always has exactly n complex roots. Some of these roots may of course be repeated, in which case there are fewer than n distinct roots. In summary, we have

Theorem 4.1 : If T is an arbitrary operator on H, then the eigenvalues of T constitute a non-empty finite subset of the complex plane. Furthermore, the number of points in this set does not exceed the dimension n of the space H.

The set of eigenvalues of T is called its spectrum, and is denoted by $\sigma(T)$. For future reference, we observe that $\sigma(T)$ is a compact subspace of the complex plane.

It should now be reasonably clear why we required in the definition of a Hilbert space that its scalars be the complex numbers. The reader will easily convince himself that in the Euclidean plane the operation of rotation about the origin through 90 degrees is an operator on this real Banach space which has no eigenvalues at all, for no non-zero vector is transformed into a real multiple of itself.

The existence of eigenvalues is therefore linked in an essential way to properties of the complex numbers which are not enjoyed by the real numbers, and the most significant of these properties is that stated in the fundamental theorem of algebra.

The mechanism of matrices and determinants turns out to be simply a device for making effective use of this theorem in our basic problem of proving that eigenvalues exist. We also remark that Theorem A and its proof remain valid in the case of an arbitrary linear transformation on any complex linear space of finite dimension $n > 0$.

Problems

1. Let T be an operator on H, and prove the following statements:
   (a) T is singular $\Leftrightarrow 0 \in \sigma(T)$;
   (b) if T is non-singular, then $\lambda \in \sigma(T) \Leftrightarrow \lambda^{-1} \in \sigma(T^{-1})$;
   (c) if A is non-singular, then $\sigma(ATA^{-1}) = \sigma(T)$;
   (d) if $\lambda \in \sigma(T)$, and if p is any polynomial, then $p(\lambda) \in \sigma\big(p(T)\big)$;
   (e) if $T^k = 0$ for some positive integer k, then $\sigma(T) = \{0\}$.
2. Let the dimension n of H be 2, let $B = \{e_1, e_2\}$ be a basis for H, and assume that the determinant of a 2×2 matrix $[\alpha_{ij}]$ is given by α_{11}.
   (a) Find the spectrum of the operator T on H defined by $Te_1 = e_2$ and $Te_2 = -e_1$.
   (b) If T is an arbitrary operator on H whose matrix relative to B is $[\alpha_{ij}]$, show that $T^2 - (\alpha_{11} + \alpha_{22})T + (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})I = 0$. Give a verbal statement of this result.

4.2 THE SPECTRAL THEOREM

We now return to the central purpose of this chapter, namely, the statement and proof of the spectral theorem.

Let T be an arbitrary operator on H. We know by Theorem 61-A that the distinct eigenvalues of T form a non-empty finite set of complex numbers. Let $\lambda_1, \lambda_2, \dots, \lambda_m$be these eigenvalues; let $M_1, M_2, \dots, M_m$be their corresponding eigenspaces; and let $P_1, P_2, \dots, P_m$be the projections on these eigenspaces. We consider the following three statements.

I. The M_i's are pairwise orthogonal and span H.

II. The P_i's are pairwise orthogonal, $I = \sum_{i=1}^{\infty} P_i$and $T = \sum_{i=1}^{\infty} \lambda_i P_i$.

III. T is normal.

We take the spectral theorem to be the assertion that these statements are all equivalent to one another. It was proved in the introduction to this chapter that I ⇒ II ⇒ III. We now complete the cycle by showing that II⇒ 1.

The hypothesis that T is normal plays its most critical role in our first theorem.

Theorem 4.2 If T is normal, then x is an eigenvector of T with eigenvalue $\lambda \Leftrightarrow x$ is an eigenvector of T* with eigenvalue $\bar{\lambda}$.

Proof.

Since T is normal, it is easy to see that the operator $T - \lambda I$ (whose adjoint is T* $-\bar{\lambda}I$) is also normal for any scalar λ. By Theorem 58-C, we have

$$\|Tx - \lambda x\| = \left\|T^* x - \bar{\lambda} x\right\|$$

for every vector x, and the statements of the theorem follow at once from this.

Theorem 4.3 If T is normal, then the M_i's are pairwise orthogonal.

Proof.

Let x_i and x_j be vectors in M_i and $M_j$for $i \neq j$, so that $Tx_i = \lambda_i x_i$and $Tx_j = \lambda_j x_j$. The preceding theorem shows that

$$\lambda_i(x_i x_j) = (\lambda_i x_i x_j)$$

$$= (T x_i x_j)$$

$$= (x_i, T^* x_j)$$

$$= (x_i, \overline{\lambda_j} x_j)$$

$$= \lambda_j (x_i x_j)$$

and since $\lambda_i \neq \lambda_j$ it is clear that we must have $(x_i x_j) = 0$.

Our next step is to prove that the M_i's span H when T is normal, and for this we need the following preliminary fact.

Theorem 4.4 : If T is normal, then each M_i reduces T.

Proof.

It is obvious that each M_i is invariant under T, so it suffices, by Theorem 59-C, to show that each M_i is also invariant under T*. This is an immediate consequence of Theorem A, for if x_i is a vector in M_i, so that $Tx_i = \lambda_i x_i$then $T^* x_i = \overline{\lambda_i} x_i$ is also in M_i.

Finally, we have

Theorem 4.5 : If T is normal, then the M_i's span H.

Proof.

The fact that the M_i's are pairwise orthogonal implies, by Theorems 59-F and 59-G, that $M = M_1 + M_2 + \cdots + M_m$is a closed linear subspace of H, and that its associated projection is

$$P = P_1 + P_2 + \cdots + P_m$$

Since each M_i reduces T, we see by Theorem 59-E that $TP_i = P_i T$ for each P_i. It follows from this that $TP = PT$, so M also reduces T, and consequently $M^\perp$ isinvariant under T. If

$M^{\perp} \neq \{0\}$, then, since all the eigenvectors of T are contained in M, the restriction of T to $M^{\perp}$ is an operator on a non-trivial finite-dimensional Hilbert space which has no eigenvectors, and hence no eigenvalues. Theorem 61-A shows that this is impossible. We therefore conclude that $M^{\perp} = \{0\}$, so M = H and the M_i's span H.

This completes the proof of the spectral theorem and, in particular, of the fact that if T is normal, then it has a spectral resolution

$$T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m \qquad (1)$$

We now make several observations which will be useful in carrying out our promise to show that this expression for T is unique. Since the P_i's are pairwise orthogonal, if we square both sides of (1) we obtain

$$T^2 = \sum_{i=1}^{m} \lambda_i^2 P_i$$

More generally, if m is any positive integer, then

$$T^n = \sum_{i=1}^{m} \lambda_i^n P_i \qquad (2)$$

If we make the customary agreement that $T^0 = I$, then the fact that $\sum_{i=1}^{m} P_i$shows that (2) is also valid for the case $n = 0$. Next, let $p(z)$ be any polynomial, with complex coefficients, in the complex variable z. By taking linear combinations, (2) can evidently be extended to

$$P(T) = \sum_{i=1}^{m} P(\lambda_i) P_i \qquad (3)$$

We would like to find a polynomial p such that the right side of (3) collapses to a specified one of the P_i's, say P_j. What is needed is a polynomial p_j with the property that $p_j(\lambda_i) = 0$if $i \neq j$ and $p_j(\lambda_j) = 1$. We define p_j as follows:

$$p_j(z) = \frac{(z-\lambda_1)\dots(z-\lambda_{j-1})(z-\lambda_{j+1})\dots(z-\lambda_m)}{(\lambda_j-\lambda_1)\dots(\lambda_j-\lambda_{j-1})(\lambda_j-\lambda_{j+1})\dots(\lambda_j-\lambda_m)}$$

Since p_j is a polynomial, and since $p_j(\lambda_i) = \delta_{ij}$, (3) yields

$$P_j = p_j(T) \qquad (4)$$

In order to interpret these remarks to our advantage, we point out that only three facts about (1) have been used in obtaining (4): the λ_i's are distinct complex numbers; the P_i's are pairwise orthogonal projections; and $I = \sum_{i=1}^{m} P_i$. By using these properties of (1), and these alone, we have shown that the P_i's are uniquely determined as specific polynomials in T.

We now assume that we have another expression for T similar to (1),

$$T = \alpha_1 Q_1 + \alpha_2 Q_2 + \cdots + \alpha_k Q_k \qquad (5)$$

and that this is also a spectral resolution of T, in the sense that the α_i's are distinct complex numbers, the Q_i's are non-zero pairwise orthogonal projections, and $I = \sum_{i=1}^{k} Q_i$. We wish to show that (5) is actually identical with (1), except for notation and order of terms. We begin by proving, in two steps, that the α_i's are precisely the eigenvalues of T. First, since $Q_i \neq 0$, there exists a non-zero vector x in the range of Q_i and since $Q_i x = x$ and $Q_i x = 0$ for $j \neq i$, we see from (5) that $Tx = \alpha_i x$, so each α_i is an eigenvalue of T. Next, if λ is an eigenvalue of T, so that $Tx = \lambda x$ for some non-zero x, then

$$Tx = \lambda x$$

$$= \lambda I x$$

$$= \lambda \sum_{i=1}^{k} Q_i \, x$$

$$= \sum_{i=1}^{k} \lambda Q_i \, x$$

$$and \qquad Tx = \sum_{i=1}^{k} \alpha_i Q_i x$$

$$so \qquad \sum_{i=1}^{k} (\lambda - \alpha_i) Q_i x = 0$$

Since the $Q_i x$'s are pairwise orthogonal, the non-zero vectors among them—there is at least one, for$x \neq 0$ are linearly independent, and this implies that $\lambda = \alpha_i$ for some i.

These arguments show that the set of α_is equals the set of λ_is, and therefore, by changing notation if necessary, we can write (5) in the form

$$T = \lambda_1 Q_1 + \lambda_2 Q_2 + \cdots + \lambda_m Q_m \qquad (6)$$

The discussion in the preceding paragraph now applies to (6) and gives

$$Q_i = p_i(T) \qquad (7)$$

for every j. On comparing (7) with (4), we see that the Q_j's equal the P_j's. This shows that (5) is exactly the same as (1) except for notation and the order of terms and completes our proof of the fact that the spectral resolution of T is unique.

We conclude with a brief look at the matrix interpretation of statements I and II at the beginning of this section.

Assume that I is true, that is, that the eigenspaces $M_1, M_2, \dots, M_m$of T are pairwise orthogonaland span H. For each M_i, choose a basis which consists of mutually orthogonal unit vectors. This can always be done, for a basis of this kind called an orthonormal basis is precisely a complete orthonormal set for M_i.

It is easy to see that the union of these little bases is an orthonormal basis for all of H; and relative to this, the matrix of T has the following diagonal form (all entries off the main diagonal are under- stood to be 0):

$$\begin{bmatrix} \lambda_1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \lambda_1 & & & & & & & \\ & & & \lambda_2 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \lambda_2 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & \lambda_m & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_m \end{bmatrix}$$

We next assume that H has an orthonormal basis relative to which the matrix of 7 is diagonal.

If we rearrange the basis vectors in such a way that equal matrix entries adjoin one another on the main diagonal, then the matrix of T relative to this new orthonormal basis will have the form (8). It is easy to see from this that T can be written in the form

$$T = \sum_{i=1}^{m} \lambda_i P_i$$

where theλ_i's are distinct complex numbers, the P_i's are non-zero pairwise orthogonal projections, and $I = \sum_{i=1}^{m} P_i$. The uniqueness of the spectral resolution now guarantees that the λ_i's are the distinct eigenvalues of T and that the P_i's are the projections on the corresponding eigenspaces. The spectral theorem tells us that statements I, II, and III are equivalent to one another. The above remarks carry us a bit further, for they constitute a proof of the fact that these statements are also equivalent to

IV. There exists an orthonormal basis for H relative to which the matrix of T is diagonal.

It is interesting to realize that the implication III ⇒ IV, which we proved by showing that III ⇒I and I⇒ IV, can be made to depend more directly on matrix computations. This proof is outlined in the last three problems below.

Problems

1. Show that an operator T on H is normal⇔ its adjoint T* is a polynomial in T.
2. Let T be an arbitrary operator on H, and N a normal operator. Show that if 7 commutes with N, then T also commutes with N*.
3. Let T be a normal operator on H with spectrum $[\lambda_1, \lambda_2, \dots, \lambda_m]$, and use the spectral resolution of T to prove the following statements:
   (a) T is self-adjoint ⇔each λ_i is real; (b) T is positive ⇔ $\lambda_i \geq 0$ for each i (c) T is unitary ⇔ $|\lambda_i| = 1$ for each i.
4. Show that a positive operator T on H has a unique positive square root; that is, show that there exists a unique positive operator A on H such that A^2 = T.

CHAPTER - 5

5.1 THE DEFINITION AND SOME EXAMPLES

A Banach algebra is a complex Banach space which is also an algebra with identity 1, and in which the multiplicative structure is related to the norm by the following requirements:

(1) $\|xy\| \leq \|x\|\|y\|$;
(2) $\|I\| = I$.

It follows from (1) that multiplication is jointly continuous in any Banach algebra, that is, that if $x_n \to x$and $y_n \to y$then $x_n y_n \to xy$

(proof:

$$\|x_n y_n - xy\| = \|x_n(y_n - y) + (x_n - x)y\|$$

$$\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|$$

A Banach subalgebra of a Banach algebra A is a closed subalgebra of A which contains 1. The Banach subalgebras of A are precisely those subsets of A which are themselves Banach algebras with respect to the same algebraic operations, the same identity, and the same norm.

The definition of a Banach algebra is sometimes given without the restriction that the scalars are the complex numbers. The complex case, however, is the only one that concerns us, and by framing the definition as we do, we avoid the necessity of treating the additional complications which arise in the real case. We have further assumed, for the sake of simplicity, that every Banach algebra has an identity. It is possible, at a considerable sacrifice of clarity, to develop most of the important ideas without this assumption, and this is done whenever the primary purpose of the theory is the study of group algebras of locally compact but not discrete groups. Since our attention will be directed chiefly to the structure of operator algebras, there is no need for us to strain for the added generality obtained by not requiring the presence of an identity.

The Banach algebras of principal interest to us are described in the following examples. The reader will notice that they all consist of functions or operators and that the linear operations in all of them are defined pointwise. They can be classified in a general way into function algebras, operator algebras, or group algebras, according as multiplication is defined pointwise, by composition, or by convolution.

Example 1. (a) One of the most important Banach algebras is the set $\mathcal{C}$(X) of all bounded continuous complex functions defined on a topological space X. The case in which X is a compact Hausdorff space will have particular significance for our later work. If X has only one point, then $\mathcal{C}$(X) can be identified with the simplest of all Banach algebras, the algebra of complex numbers.

(b) Consider the closed unit disc $D = \{z : |z| < 1\}$ in the complex plane. The subset of $\mathcal{C}$(D) which consists of all functions analytic in the General Preliminaries on Banach Algebras 303 interior of D is obviously a subalgebra which contains the identity. A simple application of Morera's theorem from complex analysis shows that it is closed and is therefore a Banach subalgebra of $\mathcal{C}$(D). This Banach algebra is called the disc algebra. It has a number of interesting proper- ties, which are, of course, intimately related to the special character of its functions.

Example 2. (a) If B is a non-trivial complex Banach space, then the set $\mathcal{C}$(B) of all operators on B isa Banach algebra. We assume that B is non-trivial in order to guarantee that the identity operator is an identity in the algebraic sense.

(b) If we consider a non-trivial Hilbert space H, then $\mathcal{C}$(H) is a Banach algebra. This is a special case of $\mathcal{C}$(B), and it is important to observe that additional structure is present here, namely, the adjoint operation T→T*.

(c) A-subalgebra of$\mathcal{C}$(H) is said to be self-adjoint if it contains the adjoint of each of its operators. Banach subalgebras of $\mathcal{C}$(H)'s which are self-adjoint are called C*-algebras. We shall return to the subject of commutative C*-algebras in Chap. 14.

(d) The weak operator topology on $\mathcal{C}$(H) is the weak topology generated by all functions of the form T→(Tx,y); that is, it is the weakest topology with respect to which all these functions are continuous. It is easy to see from the inequality$|(Tx, y) - (T_0x, y)| \leq \|T - T_0\|\|x\|\|y\|$ that this topology is weaker than the usual norm topology, so that its closed sets are also closed in the usual sense. A C'*-algebra with the further property of being closed in the weak operator topology is called a W*-algebra. Algebras of this kind are also called rings of operators, or von Neumann algebras. They are among the most interesting of all Banach algebras, but their theory is quite beyond the scope of this book.

Example 3. (a) If $G = \{g_1, g_2 \ldots, g_n\}$ is a finite group, then its group algebra $L_1(G)$ is the set of all complex functions defined on G. Addition and scalar multiplication are defined pointwise, and the norm by $\|f\| = \sum_{i=1}^{m} |f(g_i)|$. In order to see what underlies the definition of multiplication, it is convenient to regard a typical element f of $L_i(G)$ as a formal sum $\sum_{i=1}^{m} \alpha_i g_i$, where α_i is the value of f atg_i. With this interpretation, we use the given multiplication in G to define multiplication in $L_1(G)$, as follows:

$$\left(\sum_{i=1}^{n} \alpha_i g_i\right)\left(\sum_{j=1}^{n} \beta_j g_j\right) = \sum_{k=1}^{n} \gamma_k g_k \quad (1)$$

$$\gamma_k = \sum_{g_i g_j = g_k} \alpha_i \beta_j \quad (2)$$

The meaning of the sum in (2) is that the summation is to be extended over all subscripts i and j such that $g_i g_j = g_k$. In effect, therefore, we formally multiply out the sums on the left of (1), and we then gather together all the resulting terms which contain the same element of G. With these ideas as an intuitive guide, we revert to our first point of view, in which the elements of $L_1(G)$ are functions, and we see that our definition of multiplication can be expressed in the following way. If two functions f and g in $L_1(G)$ are given, then their product, which is denoted by f *g and called their convolution, is that function whose value at g_k is

$$(f * g)(g_k) = \sum_{g_i g_j = g_k} f(g_i) g(g_i)$$

$$= \sum_{k=1}^{n} f\left(g_k g_j^{-1}\right) g\left(g_j\right) \quad (3)$$

We note that if each element of G is identified with the function whose value is 1 at that element and 0 elsewhere, then G becomes a subset of $L_1(G)$. Further, multiplication in G agrees with convolution in $L_1(G)$, and the element of $L_1(G)$ which corresponds to the identity in G is an identity for $L_1(G)$. We conclude this description by observing that every element of G has norm 1, so that $||1|| = 1$, and that the basic norm inequality for a Banach algebra is satisfied:

$$\|f * g\| = \sum_{k=1}^{n} |(f * g)(g_k)|$$

$$= \sum_{k=1}^{n} \left| \sum_{j=1}^{n} f(g_k g_j^{-1}) g(g_j) \right|$$

$$\leq \sum_{k=1}^{n} \sum_{j=1}^{n} |f(g_k g_j^{-1})| |g(g_j)|$$

$$= \sum_{k=1}^{n} \sum_{j=1}^{n} |f(g_k g_j^{-1})| |g(g_j)|$$

$$= \sum_{k=1}^{n} |g(g_j)| \sum_{j=1}^{n} |f(g_k g_j^{-1})|$$

$$= \sum_{k=1}^{n} |g(g_j)| \, \|f\|$$

$$= \|f\| \sum_{k=1}^{n} |g(g_j)|$$

$$= \|f\| \|g\|$$

(b) Let $G = \{\dots, -2, -1, 0, 1, 2, \dots\}$ be the additive group of integers. Its group algebra $L_1(G)$ is the set of all complex functions f defined on G for which $\sum_{n=-\infty}^{\infty} |f(n)|$ converges. The linear operations are defined pointwise, the norm by $\|f\| = \sum_{n=-\infty}^{\infty}|f(n)|$ and the convolution of f and g—see Eq. (3)—by

$$(f * g)(n) = \sum_{n=-\infty}^{\infty} f(n-m)g(m)$$

Just as in (a), G is contained in $L_1(G)$ in a natural way, and $L_1(G)$ isa Banach algebra. Any attempt to discuss the group algebra of a non- discrete topological group like the real line must clearly be based on an adequate theory of integration. It should also have available a theory of Banach algebras in which no identity is assumed to be present. These ideas constitute a rich and beautiful field of modern analysis. They are, however, outside the scope of this work.

The Banach algebras described above are many and diverse, and there are yet others which we have not mentioned. Our attention in the following chapters will be centered on $\mathcal{C}$(X)'s and commutative C*-algebras, but the general theory we develop is equally applicable to all. It is worthy of notice that an arbitrary Banach algebra A can be regarded as a Banach subalgebra of $\mathcal{C}$(A). Ina sense, therefore, Example 2a and its Banach subalgebras include all possible Banach algebras. To see this, we recall from Problem 45-4 that a→ M_a, where $M_a(x) = ax$, isan isomorphism of A into $\mathcal{C}$(A). It is easy to see that M_1 is the identity operator on A, so all that remains is to observe that $\|a\| \leq \|M_a\|$ for every a (proof: $\|M_a(x)\| = \|ax\| \leq \|a\|\|x\|$shows that$\|M_a\| \leq \|a\|$, and the fact that $\|a\| \leq \|M_a\|$ follows from

$$\|M_a\| = \sup\{\|M_a(x)\|: \|x\| \leq 1\} \geq \|M_a(1)\| = \|a\|$$

The mapping $a \to M_a$, is thus an isometric isomorphism of A onto a Banach subalgebra of $\mathcal{C}$(A), and it allows us to identify the abstract Banach algebra A with a concrete Banach algebra of operators on A.

5.2 REGULAR AND SINGULAR ELEMENTS

Let A be a Banach algebra. We denote the set of regular elements in A by G, and its complement, the set of singular elements, by S. It is clear that G contains 1 and is a group, and that S contains 0. Several important issues depend on the character of G and S. Our first resultalong these lines is

Theorem 5.1 : Every element x for which $\|x-1\|<1$ is regular, and the inverse of such an element is given by the formula $x^{-1}=1+\sum_{n=1}^{\infty}(1-x)^n$

Proof.

If we put $r=\|x-1\|$, so that $r<1$, then

$$\|(1-x)^n\| \leq \|1-x\|^n = r^n$$

shows that the partial sums of the series $\sum_{n=1}^{\infty}(1-x)^n$ form a Cauchy sequence in A. Since A is complete, these partial sums converge to an element of A, which we denote by $\sum_{n=1}^{\infty}(1-x)^n$. If we define y by$y=\sum_{n=1}^{\infty}(1-x)^n$, then the joint continuity of multiplication in A implies that

$$y - xy = (1-x)y$$

$$= (1-x) + \sum_{n=2}^{\infty}(1-x)^n$$

$$= \sum_{n=1}^{\infty}(1-x)^n$$

$$= y - 1$$

So$xy = 1$. Similarly, $yx = 1$.

We now use this as a tool to prove

Theorem 5.2 : G is an open set, and therefore S is a closed set.

Proof.

Let x_0 be an element in G, and let x be any element in A such that$\|x - x_0\| < 1/\|x_0^{-1}\|$. It is clear that

$$\|x_0^{-1}x - 1\| = \|x_0^{-1}(x - x_0)\|$$

$$\leq \|x_0^{-1}\|\|x - x_0\|$$

$$< 1$$

so we see by Theorem A that$x_0^{-1}x$ is in G. Since $x = x_0(x_0^{-1}x)$, it follows that x is also in G, so G is open.

It was shown that every Banach space is locally connected, so A is also locally connected.

Theorem 5.3 : The mapping $x \to x^{-1}$ of G into G is continuous and is therefore a homeomorphism of G onto itself.

Proof.

Let x_0 be an element of G, and x another element of G such that$\|x - x_0\| < 1/(\|x_0^{-1}\|)$. Since

$$\|x_0^{-1}x - 1\| = \|x_0^{-1}(x - x_0)\|$$

$$\leq \|x_0^{-1}\|\|x - x_0\|$$

$$< \frac{1}{2}$$

we see by Theorem A that $x_0^{-1}x$ is in G and

$$x_0^{-1}x = (x_0^{-1}x)^{-1}$$

$$= 1 + \sum_{n=1}^{\infty}(1 - x_0^{-1}x)^n$$

Our conclusion now follows from

$$\|x^{-1} - x_0^{-1}\| = \|(x^{-1} x_0 - 1)x_0^{-1}\|$$

$$\leq \|x_0^{-1}\|\|x^{-1} x_0 - 1\|$$

$$= \|x_0^{-1}\| \left\| \sum_{n=1}^{\infty} (1 - x_0^{-1}x)^n \right\|$$

$$\leq \|x_0^{-1}\| \sum_{n=1}^{\infty} \|1 - x_0^{-1}x\|^n$$

$$= \|x_0^{-1}\|\|1 - x_0^{-1}x\| \sum_{n=0}^{\infty} \|1 - x_0^{-1}x\|^n$$

$$= \frac{\|x_0^{-1}\|\|1 - x_0^{-1}x\|}{1 - \|1 - x_0^{-1}x\|}$$

$$< 2\|x_0^{-1}\|\|1 - x_0^{-1}x\|$$

$$\leq 2\|x_0^{-1}\|\|x - x_0\|$$

If x is an element in A, it should always be kept in mind that the regularity or singularity of x depends on A as well as on x itself. If x is regular in A, and if we pass to a Banach subalgebra A' of A which also contains x, then x may lose its inverse and become singular in A'. By the same token, if x is singular in A, and if A is regarded as a Banach subalgebra of a larger Banach algebra A'', then x may acquire an inverse and become regular in A''. In the next section, we study certain elements in A which are singular and remain singular with respect to all possible enlargements of A.

5.3 TOPOLOGICAL DIVISORS OF ZERO

An element z in our Banach algebra A is called a topological divisor of zero if there exists a sequence $\{z_n\}$ in A such that $\|z_n\| = 1$ and either $zz_n \to 0$ or $z_n z \to 0$. Itis clear that every divisor of zero is also a topological divisor of zero. We denote the set of all topological divisors of zero by Z.

Theorem 5.4 Z is a subset of S.

Proof.

Let z be an element of Z and $\{z_n\}$ a sequence such that $\|z_n\| = 1$ and (say) $zz_n \to 0$. If z were in G, then by the joint continuity of multi- plication we would have $z^{-1}(zz_n) = z_n \to 0$, contrary to $\|z_n\| = 1$.

Our next theorem relates to the manner in which Z is distributed within S.

Theorem 5.5 : The boundary of S is a subset of Z.

Proof.

Since S is closed, its boundary consists of all points in S which are limits of convergent sequences in G.

We show that if z is such a point, that is, if z is in S and there exists a sequence $\{r_n\}$ in G such that $r_n \to z$, then z is in Z.

First, we see from $r_n^{-1}z - 1 = r_n^{-1}(z - r_n)$ that the sequence $\{r_n^{-1}\}$ is unbounded; for otherwise, we would have

$$\|r_n^{-1}z - 1\| < 1$$

for some n, so that $r_n^{-1}z$, and therefore$z = r_n(r_n^{-1}z)$, would be regular. Since $\{r_n^{-1}\}$ is unbounded, we may assume that$\|r_n^{-1}\| \to \infty$.

If z_n, is now defined by$z_n = r_n^{-1}/\|r_n^{-1}\|$, then our conclusion follows from the observations that $\|z_n\| = 1$ and

$$zz_n = \frac{zr_n^{-1}}{\|r_n^{-1}\|}$$

$$= \frac{1 + (z - r_n)r_n^{-1}}{\|r_n^{-1}\|}$$

$$= \frac{1}{\|r_n^{-1}\|} + (z - r_n)z_n \to 0$$

In order to understand the significance of these facts, let us suppose that A is imbedded as a Banach subalgebra in a larger Banach algebra A'.

As we remarked in the previous section, an element which is singular in A may cease to be so in A'.

However, if it is a topological divisor of zero in A, then it is in A' as well, so it is singular in A'. The topological divisors of zero in A are thus “permanently singular,” in the sense that they are singular and remain so with respect to every possible enlargement of the containing Banach algebra. Theorem B tells us that no matter what happens to S as a whole in such a process, its boundary is “permanent” in this sense.

5.4 THE SPECTRUM

Let T be an operator on a non-trivial Hilbert space. In the previous chapter, we defined the spectrum of T' to be the set

$$\sigma(T) = \{\lambda : T - \lambda I \ is \ singular\}$$

and we devoted a good deal of attention to the geometric ideas leading to this concept. We found—at least in the finite-dimensional case—that a number in $\sigma(T)$ is a value assumed by T, in the sense that T acts on some non-zero vector as if it were scalar multiplication by that number. We shall see later that this formulation of the meaning of the spectrum has a much wider significance than we might at first suspect.

Let us now consider an element z in our general Banach algebra A. By analogy with the above, we define the spectrum of x to be the following subset of the complex plane:

$$\sigma(x) = \{\lambda : x - \lambda I \ is \ singular\}$$

Whenever it is desirable to express the fact that the spectrum of x depends on A as well as x, we use the notation $\sigma_A(x)$. It is easy to see that $x - \lambda I$ is a continuous function of $\lambda$with values in A; and since the set of singular elements in A is closed, it follows at once that $\sigma(x)$ is closed. We further observe that $\sigma(x)$ is a subset of the closed disc$\{x: |x| < ||x||\}$, for if λisa complex number such that$|\lambda| \geq \|x\|$, then $\|x/\lambda\| < 1, \|1 - (1 - x/\lambda)\| < 1, 1 - x/\lambda$is regular, and therefore $x - \lambda I$ is regular.

Our first task is to establish the fact that $\sigma(x)$ is always non-empty, and for this we need a few preliminary notions. The resolvent set of x, denoted by $\rho(x)$, is the complement of $\sigma(x)$; it is clearly an open subset of the complex plane which contains $\{z: |z| \ > \ ||z||\}$. The resolvent of x is the function with values in A defined on $\rho(x)$ by

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

tells us that $x(\lambda)$ is a continuous function of λ and the fact that $x(\lambda) = \lambda^{-1}\left(\frac{x}{\lambda} - 1\right)^{-1}$ for $\lambda =$ 0implies that $x(\lambda) \to 0$ as $\lambda \to \infty$. If λ and μ are both in$\rho(x)$, then

$$x(\lambda) = x(\lambda)[x - \mu I]x(\mu)$$

$$= x(\lambda)[x - \lambda I + (\lambda - \mu)I]x(\mu)$$

$$= [I + (\lambda - \mu)x(\lambda)]x(\mu)$$

$$= x(\mu) + (\lambda - \mu)x(\lambda)x(\mu)$$

$$x(\lambda) - x(\mu) = (\lambda - \mu)x(\lambda)x(\mu)$$

This relation is called the resolvent equation.

Theorem 5.6 : $\sigma(x)$ is non-empty.

Proof.

Let f be a functional on A—that is, an element of the conjugate space A*—and define $f(\lambda)$ by $f(\lambda) = f\big(x(\lambda)\big)$. It is clear that $f(\lambda)$ is a complex function which is defined and continuous on the resolvent set $\rho(x)$. The resolvent equation shows that

$$\frac{f(\lambda) - f(\mu)}{\lambda - \mu} = f\big(x(\lambda)x(\mu)\big)$$

and it follows from this that

$$\lim_{\lambda \to \mu} \frac{f(\lambda) - f(\mu)}{\lambda - \mu} = f(x(\mu)^2)$$

so $f(\lambda)$ has a derivative at each point of $\rho(x)$. Further,

$$|f(\lambda)| \leq \|f\| \|x(\lambda)\|$$

So $f(\lambda) \to 0$ as $\lambda \to \infty$.

We now assume that $\sigma(x)$ is empty, so that $\rho(x)$ is the entire complex plane. Liouville's theorem from complex analysis allows us to conclude that $f(\lambda) = 0$ for all λ. Since f is an arbitrary functional on A, implies that $x(\lambda) = 0$ forall λ.

This is impossible, for no inverse can equal 0, and therefore it cannot be true that $\sigma(x)$ is empty.

If the reader is surprised by the appearance of Liouville's theorem in such a context, he should recall two facts. First, our proof is a special case of the above result, required the use of the fundamental theorem of algebra. And second, the fundamental theorem of algebra is most commonly proved as a simple consequence of Liouville's theorem. It is therefore only to be expected that some tool from analysis comparable in depth with Liouville's theorem should be necessary for the proof of Theorem A.

Now that we know that $\sigma(x)$ is non-empty, we also know that it is a compact subspace of the complex plane. The number $r(x)$ defined by

$$r(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}$$

is called the spectral radius of x. It is clear that$0 \leq r(x) \leq \|x\|$. The concept of the spectral radius will be useful in certain parts of our later work.

We recall that a division algebra is an algebra with identity in which each non-zero element is regular. The most important single consequence of Theorem A is

Theorem 5.7 : If A is a division algebra, then it equals the set of all scalar multiples of the identity.

Proof.

We must show that if x is an element of A, then x equals λI for some scalar λ. Suppose, on the contrary, that $x \neq \lambda I$ for every λ.

Then $x = \lambda I \neq 0$for every λ, $x - \lambda I$ is regular for every λ, and therefore $\sigma(x)$ isempty. This contradicts Theorem 5.6 and completes the proof.

The mapping $\lambda I \rightarrow \lambda$ is clearly an isometric isomorphism of the set of all scalar multiples of the identity onto the Banach algebra C of all complex numbers.

We may therefore identify this set with C; and in terms of this identification, any Banach algebra which is a division algebra equals C. This fact is the foundation on which we build the structure theory presented in the next chapter.

It is obvious that C itself, which is the simplest of all Banach algebras, is a division algebra, so Theorem 5.7 characterizes C as the only Banach algebra with this property. In the next two theorems, we give some other interesting characterizations of C among all possible Banach algebras.

Since 0 is a divisor of zero, it is a topological divisor of zero in every Banach algebra. In the Banach algebra C, 0 is plainly the only topological divisor of zero. Conversely, we have

Theorem 5.8 : If 0 is the only topological divisor of zero in A, then A = C.

Proof.

Let x be an element of A. Its spectrum $\sigma(x)$ is non-empty, so it has a boundary point A; and $x - \lambda I$ is easily seen to be a boundary point of the set S of all singular elements.

By Theorem, $x - \lambda I$ is a topological divisor of zero, so it follows from our hypothesis that$x - \lambda I = 0$ or $x = \lambda I$.

The basic link between multiplication in A and the norm is given by the inequality $\|xy\| \leq \|x\|\|y\|$, and when A = C, this inequality can be reversed.

The following result shows to what extent this reversibility is true in general.

Theorem 5.9 : If the norm in A satisfies the inequality $\|xy\| \geq K\|x\|\|y\|$ for some positive constant K, then A = C.

Proof.

In the light of Theorem, it suffices to observe that the hypothesis here implies that 0 is the only topological divisor of zero.

We next look into the question of what happens to the spectrum of an element x in A when A is enlarged.

Theorem 5.10 : If A is a Banach subalgebra of a Banach algebra A', then the spectra of an element $x$in A with respect to A and A' are related as follows: (1) $\sigma_{A'}(x) \subseteq \sigma_A(x)$; (2) each boundary point of $\sigma_A(x)$ is also a boundary point of $\sigma_{A'}(x)$.

Proof.

If $x - \lambda I$ is singular in A', then it is certainly singular in A, so (1) is clear. To prove (2), we let λ be a boundary point of $\sigma_A(x)$.

It is easy to see that $x - \lambda I$ is a boundary point of the set of singular elements in A, it is a topological divisor of zero in A.

It is therefore a topological divisor of zero in A' as well, so it is singular in A' and λ is in $\sigma_{A'}(x)$. The fact that A is actually a boundary point of $\sigma_{A'}(x)$ is immediate from (1), so the proof of (2) is complete.

This result shows that in general the spectrum of an element shrinks when its containing Banach algebra is enlarged, and further, that since its boundary points cannot be lost in this process, it must shrink by "hollowing out." An illuminating example of this phenomenon is provided by the disc algebra A of all complex functions which are defined and continuous on $D = \{z: |z| \leq 1\}$ and analytic in the interior of this set. If $f$is a function in A, then the maximum modulus theorem from complex analysis implies that

$$\|f\| = sup\,\{|f(z)|: |z| \leq 1\}$$

$$= sup\,\{|f(z)|: |z| = 1\}.$$

This allows us to identify A with the Banach algebra of all the restrictions of its functions to the boundary of D, which is a Banach subalgebra of $A' = \mathcal{C}(\{z: |z| = 1\})$. If we now consider the element f in A defined by $f(z) = z$, then it is easy to see that $\sigma_A(f)$ equals D and that $\sigma_{A'}(f)$equals the boundary of D.

5.5 THE FORMULA FOR THE SPECTRAL RADIUS

Let x be an element in our general Banach algebra A, and consider its spectral radius r(x), which is defined by

$$r(x) = sup\,\{|\lambda|: \lambda \in \sigma_A(x)\}.$$

Now let A' be the Banach subalgebra of A generated by x, that is, the closure of the set of all polynomials in x. Theorem 67-E shows that $r(x)$ has the same value if it is computed with respect to A':

$$r(x) = sup\,\{|\lambda|: \lambda \in \sigma_{A'}(x)\}.$$

This suggests quite strongly that $r(x)$ depends only on the sequence of powers of x. The formula for $r(x)$ is given in Theorem A below, and our purpose in this section is to prove it. It is convenient to begin with the following preliminary result.

Lemma : $\sigma(x^n) = \sigma(x)^n$.

Proof.

Let λ be a non-zero complex number and $\lambda_1, \lambda_2, \dots, \lambda_n$ its distinct nth roots, so that

$$x^n - \lambda I = (x - \lambda_1 I)(x - \lambda_2 I) \dots (x - \lambda_n I)$$

The statement of the lemma follows easily from the fact that $x^n - \lambda I$ is singular $\Leftrightarrow$ $x - \lambda_i I$ is singular for at least one i.

Theorem 5.11 : $r(x) = \lim \|x^n\|^{\frac{1}{n}}$.

Proof.

Our lemma shows that $r(x^n) = r(x)^n$, and since $r(x^n) \leq \|x^n\|$, we have $r(x)^n \leq \|x\|^n$ or $r(x) \leq \|x^n\|^{\frac{1}{n}}$ for every n. To conclude the proof, it suffices to show that if a is any real number such that $r(x) < a$, then $\|x^n\|^{\frac{1}{n}} \leq a$ for all but a finite number of n's, and this we now do.

If $|\lambda| = \|x\|$, then

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

$$= \lambda^{-1}\left(\frac{x}{\lambda} - 1\right)^{-1}$$

$$= -\lambda^{-1}\left(1 - \frac{x}{\lambda}\right)^{-1}$$

$$= -\lambda^{-1}\left[\sum_{n=1}^{\infty}\frac{x^n}{\lambda^n}\right] \quad (1)$$

If f is any functional on A, then (1) yields

$$f(x(\lambda)) = -\lambda^{-1}\left[f(1) + \sum_{n=1}^{\infty} f\left(\frac{x^n}{\lambda^n}\right)\right]$$

$$= -\lambda^{-1}\left[f(1) + \sum_{n=1}^{\infty} f(x^n)\lambda^{-n}\right] \quad (2)$$

for all $|\lambda| > \|x\|$. Therefore, $f(x(\lambda))$ is an analytic function in the region $|\lambda| > r(x)$ and since (2) is its Laurent expansion for $|\lambda| > \|x\|$, we know from complex analysis that this expansion is valid for $|\lambda| > r(x)$. If we now let a be any real number such that $r(x) < \alpha < a$, then it follows from the preceding remark that the series $\sum_{n=1}^{\infty} f(x^n/\alpha^n)$ converges, so its terms form a bounded sequence.

Since this is true for every f in A*, this shows that the elements x^n/α^n form a bounded sequence in A. Thus

$$\|x^n/\alpha^n\| \leq K$$

or $\|x^n\| \leq K^{\frac{1}{n}}\alpha$ for some positive constant K and every n. Since $K^{\frac{1}{n}}\alpha < a$ for every sufficiently large n, we have $\|x^n\|^{\frac{1}{n}} \leq a$ for all but a finite number of n's, and the proof is complete.

5.6 THE RADICAL AND SEMI-SIMPLICITY

Our final preliminary task is to reach a clear understanding of what is meant by the statement that our Banach algebra A is semi-simple. For this, it is necessary to give an adequate definition of the radical of A, and this in turn depends on a detailed analysis of its ideals.

We recall that an deal in A was defined in Sec. 45 to be a subset J with the following three properties:

(1) I is a linear subspace of A;

(2)$i \in I \Rightarrow xi \in I$ for every element $x \in A$;

(3) $i \in I \Rightarrow ix \in I$ for every element $x \in A$.

If I is assumed only to satisfy conditions (1) and (2) [or conditions (1) and (3)], it is called a left ideal (or a right ideal). Yor the sake of clarity and emphasis, an ideal in our previous sense—one which satisfies all three of these conditions—is often called a two-sided ideal. In the commutative case, of course, these three concepts coincide with one another.

The properties of the ideals in A are closely related to the properties of its regular and singular elements. In our work so far, the statement that an element z in A is regular has meant that there exists an element y such that $xy = yx = 1$. For our present purposes, it is useful to refine this notion slightly, as follows. We say that x is left regular if there exists an element y such that $yx = 1$; and if x is not left regular, it is called left singular. The terms right regular and right singular are defined similarly. If x is both left regular and right regular, so that there exist elements y and z such that $yx = 1$ and $xz = 1$, then the relation

$$y = y1 = y(zz) = (yx)z = 1z = z$$

shows that x is regular in the ordinary sense and that x^{-1} = y = z.

The concept of maximality for two-sided ideals was introduced in Sec. 41. By analogy, we define a maximal left ideal in A to be a proper left ideal which is not properly contained in any other proper left ideal. A straightforward application of Zorn's lemma shows that any proper left: ideal can be imbedded in a maximal left ideal; and since the zero ideal $\{0\}$ is a proper left ideal, maximal left ideals certainly exist. We now define the radical R of A to be the intersection of all its maximal left ideals. It will be convenient to abbreviate this definition by writing $R =\cap MLI$. R is clearly a proper left ideal.

These ideas can be formulated just as easily for right ideals as for left ideals, and there is no reason for giving preference to either side over the other. The purpose of the following chain of lemmas is to show that B is also the intersection of all the maximal right ideals in A, that is, that$R =\cap MRI$.

Lemma : If r is an element of R, then $1 - r$ is left regular.

Proof.

We assume that $1 - r$ is left singular, so that

$$L = A(I—r) = \{x - xr : x \in A\}$$

is a proper left ideal which contains $I - r$. We next imbed L in a maximal left ideal 7, which of course also contains $I - r$. Since r is in R, it is also in M, and therefore $I = (I - r) + r$isin M. This implies that M = A, which is a contradiction.

Lemma : If r is an element of R, then $I - r$ is regular.

Proof.

By the lemma just proved, there exists an element s such that $s(1 — r) = 1$, so s is right regular and $s = 1 - (-s)r$. The fact that R is a left ideal implies that $(—s)r$ is in R along with r, and another application of the preceding lemma shows that $1-(-s)r = s$ is left regular. Since s is both left regular and right regular, it is regular with inverse $I-r$, so $I-r$ is also regular.

Lemma : If r is an element of R, then $I-xr$ is regular for every x.

Proof.

R is a left ideal, so xr is in R and the statement follows from the lemma just proved.

Lemma : If r is an element of A with the property that $I-xr$ is regular for every x, then r is in R.

Proof.

We assume that r is not in R, so that r is not in some maxi- mal left ideal M. It is easy to see that the set

$$M+Ar=\{m+zr\colon m\in M\ and\ x\in A\}$$

is a left ideal which contains both M and$r$, so $M + Ar = A$and

$$m+xr=1$$

for some mand x. It now follows that $1 — zr = m$ is a regular element in M, and this is impossible, for no proper ideal can contain any regular element.

The effect of these lemmas is to establish the equality of two sets:

$$\cap\, MLI = \{r\colon 1-xr\ is\ regular\ for\ every\ x\} \qquad (1)$$

Precisely the same arguments, when applied to maximal right ideals, show that

$$\cap\, MRI = \{r\colon 1-rx\ is\ regular\ for\ every\ x\} \qquad (2)$$

We now prove that all four of these sets are the same by showing that the two sets on the right of (1) and (2) are equal to one another. By sym- metry, it evidently suffices to prove the

Lemma : If $I - xr$is regular, then $I - rx$ is also regular.

Proof.

We assume that $I - xr$ is regular with inverse

$$s = (I - xr)^{-1}$$

This means, of course, that $(1 — xr)s = s(1— xr) = 1$. We leave it to the reader to show, by a simple computation, that

$$(I - rx)(I + rsx) = (I + rsx)(I - rx) = 1$$

so that $I - rx$isregular with inverse $1 + rsx$. (The formula for $(1 — rz)^{-1}$ is less mysterious than it looks, as the reader can see by inspecting the meaningless but suggestive expressions

$$s = (I - xr)^{-1} = 1 + xr + (xr)^2 + \cdots$$

and

$$(I - rx)^{-1} = 1 + rx + (rx)^2 + (rx)^3 + \cdots$$

$$= 1 + rx + rxrx + rxrxrx + \cdots$$

$$= 1 + r(1 + xr + xrxr + \cdots)x$$

$$= 1 + rsx$$

We summarize our results in

Theorem 5.12 : The radical R of A equals each of the four sets in (1) and (2) and ts therefore a proper two-sided ideal.

A is said to be semi-simple if its radical equals the zero ideal $\{0\}$, that is, if each non-zero element of A is outside of some maximal left ideal.

It will be observed that the ideas discussed above are purely algebraic in nature. They can be applied not only to our Banach algebra A, but also to any algebra or ring with identity.

Our interest, however, is in A, notably, the fact that the set S of all singular elements in A is closed.

We begin by noting that if J is any ideal in A (left, right, or two- sided), then by the joint continuity of the algebraic operations, its closure [is an ideal of the same kind. Next, since any proper ideal is contained in the proper closed set S, the closure of any proper ideal is a proper ideal of the same kind. It is an easy step from these facts to

Theorem 5.13 : Every maximal left ideal in A is closed.

Proof.

If any maximal left ideal L is not closed, then L is a proper subset of the proper left ideal L; and this cannot happen, for it contradicts the maximality of L.

Taken together, the above two theorems yield.

Theorem 5.14 The radical R of A is a proper closed two-sided ideal.

Theorem 5.15 : If I is a proper closed two-sided ideal in A, then the quotient algebra $A/I$is a Banach algebra.

Proof.

We have, A/I is a non-trivial complex Banach space with respect to the norm defined by

$$\|x + I\| = \inf\{\|x + i\|: i \in I\}$$

Further, A/I is clearly an algebra with identity $1 + I$, and

$$\|1 + I\| = \inf\{\|1 + i\|: i \in I\} \leq \|I\| = 1$$

The multiplicative inequality for the norm is easily proved as follows:

$$\|(x+I)(y+I)\| = \|xy+I\|$$

$$= \inf\{\|xy+i\| : i \in I\}$$

$$\leq \inf\{\|(x+i_1)(y+i_2)\| : i_1, i_2 \in I\}$$

$$\leq \inf\{\|x+i_1\|\|y+i_2\| : i_1, i_2 \in I\}$$

$$\leq \inf\{\|x+i_1\| : i_1 \in I\} \inf\{\|y+i_2\| : i_2 \in I\}$$

$$= \|x+I\|\|y+I\|$$

All that remains is to show that $\|1+I\| = 1$; and since we already have $\|1+I\| \leq 1$, this is an immediate consequence of the fact that $\|1+I\| = \|(1+I)^2\| \leq \|1+I\|^2$ implies $1 \leq \|1+I\|$.

Theorem 5.16 A/R is a semi-simple Banach algebra.

Proof.

It suffices to observe that the natural homomorphism $x \to x + R$ of A onto A/R induces a one-to-one correspondence between the maximal left ideals in A and those in A/R.

Study Learning Material Prepared by

Dr. S.N. LEENA NELSON M.Sc., M.Phil., Ph.D.

Associate Professor & Head, Department of Mathematics,

Women's Christian College, Nagercoil – 629 001,

Kanyakumari District, Tamilnadu, India.