Spectral Theory
of
Operators on Hilbert Space

A. D. Andrew
W. L. Green

January 2002

School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160

© 2002  A. D. Andrew and W. L. Green
Contents

1. Introduction

Part I. Introduction to Hilbert Spaces

2. Geometry of Hilbert Spaces
3. Examples of Linear Operators
4. Operators and their Adjoints
5. Classes of Operators
6. Invertibility and the Spectrum of an Operator

Part II. Banach Algebras

7. Normed algebras and the Gelfand-Mazur Theorem
8. Ideals in Banach Algebras
9. An Example - Convolution on $l_1(\mathbb{Z})$
10. The Gelfand Transform
11. Involutions and *-algebras
12. The spectral radius formula and the Gelfand-Naimark theorem

Part III. The Spectral Theorem. A Continuous Functional Calculus

13. The Spectral Theorem
14. Some Applications of the Spectral Theorem
15. Another version of the Spectral Theorem

Part IV. The Spectral Theorem. A Functional Calculus for Bounded Measurable Functions

16. Resolutions of the Identity
17. Representations and Resolutions of the Identity
18. Some Applications of the Spectral Theorem
19. More Applications of the Spectral Theorem, Compact Operators

Part V. Unbounded Operators

20. Unbounded Operators, Terminology and Examples
21. Symmetric and Self-adjoint operators
22. The Cayley Transform
23. Integration of (unbounded) measurable functions
24. The Spectral Theorem for Self-adjoint Operators
1. Introduction

These notes are an introduction to the spectral theory of operators on a Hilbert space. To begin with, you may regard spectral theory as extension of the diagonalization of a matrix.

The diagonalization of a symmetric (or normal) matrix may be given several interpretations. To say that there is a diagonal matrix \( D \) and an orthogonal matrix \( U \) with \( A = UDU^T \), is to say that the columns of \( U \) are an orthonormal basis for \( \mathbb{C}^n \) consisting of eigenvectors for \( A \). If we group together the columns corresponding to the distinct eigenvalues, this result states that \( A = \sum \lambda_i P_i \), where \( P_i \) is the orthogonal projection onto the eigenspace for \( \lambda_i \). In these notes we extend this type of representation, first to certain classes of bounded linear operators on a Hilbert space, and then to self-adjoint unbounded operators. The sums of projections will be replaced by integrals with respect to projection valued measures.

The spectral theory of operators on a Hilbert space is a rich, beautiful, and important theory. We hope you’ll see how it all holds together and how it depends on much that you have seen in other areas, including the theory of measure and integration and the theory of analytic functions.

Our approach to the Spectral Theorem will be by way of the study of \( C^* \)-algebras. Accordingly, these notes are organized in the following manner. Part I is an introduction and review of the basic geometric properties of Hilbert space and the linear operators on Hilbert space. Part II is devoted to Banach algebras and the Gelfand theory of commutative Banach algebras. In Part III this theory is applied to the algebra generated by a normal operator on a Hilbert space, and we obtain the spectral theorem with a continuous functional calculus. Part IV is devoted to extending the functional calculus to bounded measurable functions, and in Part V we develop the spectral theory of unbounded self-adjoint operators by means of the Cayley transform.
Part I. Introduction to Hilbert Spaces

2. Geometry of Hilbert Spaces

We shall assume that you are familiar with the basic geometric properties of Hilbert spaces, although we will review them briefly in this section. We'll soon see the need for using complex scalars, so we assume throughout that all vector spaces are over the complex field $\mathbb{C}$.

An inner-product space is a complex vector space $X$ equipped with a function $\langle \cdot , \cdot \rangle : X \oplus X \to \mathbb{C}$, called the inner-product, which satisfies

a. $\langle \alpha x + \beta y , z \rangle = \alpha \langle x , z \rangle + \beta \langle y , z \rangle \quad \forall x, y, z \in X, \forall \alpha, \beta \in \mathbb{C}$,

b. $\langle y , x \rangle = \overline{\langle x , y \rangle} \quad \forall x, y \in X$, (hence $\langle \cdot , \cdot \rangle$ is conjugate-linear in its second variable), and

c. $\langle x , x \rangle \geq 0$ for each $x \in X$, with $\langle x, x \rangle = 0$ if and only if $x = 0$.

One then sees (the Cauchy-Schwarz Inequality) that for every $x, y \in X$,

$|\langle x , y \rangle| \leq \langle x , x \rangle^{1/2} \langle y , y \rangle^{1/2}$ with equality if and only if $x$ and $y$ are linearly dependent. From this it follows that $\| x \| = \langle x , x \rangle^{1/2}$ defines a norm on $X$. A Hilbert space is an inner-product space which is complete under this norm.

It follows readily from the Cauchy-Schwarz Inequality that if $\{x_n\}$ converges to $x$ in $X$, and if $\{y_n\}$ converges to $y$ in $X$, then $\langle x_n , y_n \rangle$ converges to $\langle x , y \rangle$ in $\mathbb{C}$. (This result does not make use of the completeness of $X$.)

We will draw our examples mainly from three Hilbert spaces. The first is $l_2$, the space of all square summable complex sequences $\{a_i\}$ with

the inner product $\langle \{a_i\} , \{b_i\} \rangle = \sum_{i=1}^{\cdot} a_i \overline{b_i}$. We will sometimes denote this space by $l_2(\mathbb{N})$ to distinguish it from $l_2(\mathbb{Z})$, which is the space of all doubly infinite square summable complex sequences with the inner product
given by \( \langle \{a_i\}, \{b_i\} \rangle = \sum_{i=-\cdot} a_i \overline{b_i} \). In both of these spaces we will use the notation \( e_n \) for the vector which is 1 in position \( n \) and 0 elsewhere, and will usually write \( x = \sum a_n e_n \) rather than \( \{a_n\} \). The third space is \( L^2[0,1] \), the space of all (equivalence classes of) square integrable Lebesgue measurable functions on the unit interval \([0,1]\), with inner product \( \langle f, g \rangle = \int f \overline{g} \).

The geometry of Hilbert space is determined by its inner product, and is much the same as that of \( \mathbb{C}^n \). For example, an easy calculation shows that the parallelogram identity holds in such a space:

\[
\| x + y \|^2 + \| x - y \|^2 = 2 (\| x \|^2 + \| y \|^2 ) \text{ for all } x, y \in H.
\]

A more interesting (and more difficult) argument shows that if \( H \) is a Banach space in which this identity holds for all \( x \) and \( y \), then there is an inner product on \( H \) which makes it into a Hilbert space and induces the original norm.

As in the case of \( \mathbb{C}^n \) or \( \mathbb{R}^n \), we may use the inner product to define orthogonality by saying \( x \perp y \) if and only if \( \langle x, y \rangle = 0 \). If \( M \) is a subset of \( H \), the orthogonal complement of \( M \) is the subspace

\[
M^\perp = \{ x : \langle m, x \rangle = 0 \text{ } \forall \text{ } m \in M \}.
\]

By the Cauchy-Schwarz Inequality \( M^\perp \) is always a closed subspace of \( H \).

The closed convex subsets and subspaces of a general Hilbert space have many of the important properties of their analogs in \( \mathbb{C}^n \). In particular we have the following fundamental result.

**Theorem (Closest Point Theorem)** Let \( C \) be a non-empty closed convex set in a Hilbert space \( H \), and let \( x_0 \in H \). There exists a unique point \( x_c \in C \) such that

\[
\| x_0 - x_c \| = \text{dist}(x_0, C) = \inf\{ \| x_0 - x \| : x \in C \}.
\]
Moreover, \( \text{Re} \left< x_c, x_0 - x_c \right> \cdot \text{Re} \left< x, x_0 - x_c \right> \) for all \( x \) in \( C \).

The existence and uniqueness of \( x_c \) may be proved by considering a sequence \( \{x_n\} \) from \( C \) for which \( \| x_0 - x_n \| \) converges to \( \text{dist}(x_0, C) \), and showing that \( \{x_n\} \) is Cauchy. This proof is left as an exercise. However, one should have a geometric understanding of this theorem in terms of angles and orthogonality. The picture one should have of this theorem is

where the angle at \( x_c \) is obtuse. Namely, \( x_c \) is the closest point if and only if for any \( x \in C \) and any (fairly small) \( t > 0 \), we have (with \( u = x - x_c \))

\[
\| x_0 - (x_c + t u) \|^2 \cdot \| x_0 - x_c \|^2.
\]

Now

\[
\| x - (x_c + t u) \|^2 = \left< x_0 - (x_c + t u), x_0 - (x_c + t u) \right> = \| x_0 - x_c \|^2 + t^2 \| u \|^2 - 2t \text{Re} \left< x_0 - x_c, u \right>.
\]

Thus \( x_c \) is a closest point of \( C \) to \( x_0 \) if and only if \( \text{Re} \left< x_0 - x_c, x - x_c \right> \cdot 0 \) whenever \( x \in C \). If \( C \) is a subspace of \( H \), then whenever \( x \in C \), both \( x = x_c + u \) and \( y = x_c - u \) belong to \( C \), from which it follows that \( (x - x_c) \) is perpendicular to \( (x_0 - x_c) \).
A few related observations are summarized in the next few results, the proofs of which are again left to the reader.

**Proposition**. Let $M$ be a subset of a Hilbert space $H$. Then $M^\perp$ is always a closed subspace of $H$, and if $M$ is a subspace of $H$, then $M^{\perp\perp} = M$.

Moreover, $M^{\perp\perp} = M$ if and only if $M$ is closed.

**Theorem** Let $M$ be a closed subspace of a Hilbert space $H$. Any $x \in H$ may be written uniquely as $x = m + m^\perp$ with $m \in M$ and $m^\perp \in M^\perp$. The point $m$ is the closest point of $M$ to $x$. The map $P: x \to m$ is linear, idempotent ($P^2 = P$) and of norm 1.

This theorem says that if $M$ is a closed subspace of $H$, then $H$ is the direct sum of the subspaces $M$ and $M^\perp$. We say that $H$ is the **direct sum** of closed subspaces $Y$ and $Z$ if each $x \in H$ may be written uniquely as $x = y + z$, with $y \in Y$ and $z \in Z$. You can check that the assignment from $x$ to $y$ (or to $z$) is always linear, idempotent, and bounded. If $H$ and $K$ are two Hilbert spaces we can define their (external) direct sum as the linear space of all ordered pairs $(h,k)$ with $h \in H$ and $k \in K$. The direct sum $H \oplus K$ is a Hilbert space when it is given the inner product

$$\langle (h_1, k_1), (h_2, k_2) \rangle = \langle h_1, h_2 \rangle_H + \langle k_1, k_2 \rangle_K.$$

If $H$ and $K$ are normed spaces, a function $T: H \to K$ is called a **linear operator** if $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$ for all $x,y \in H$ and all $\alpha, \beta$ in $\mathbb{C}$. In the case that $K = \mathbb{C}$, we call $T$ a **linear functional**. We say that $T$ is **bounded** if the quantity

$$||T|| = \sup\{||Tx|| : ||x|| = 1\} = \sup\{\frac{||Tx||}{||x||} : x \neq 0\}$$

is finite. This is called the **norm** of the operator $T$, and it is easily seen that it is in fact a norm on the space of all bounded linear operators from $H$ to $K$, which we will denote by $\mathcal{B}(H,K)$. We use the notation $\mathcal{B}(H)$ for the space of all bounded linear operators from $H$ to itself. The range of an operator $T$ will be denoted by $\mathcal{R}(T)$. It is not hard to see that a linear operator is continuous if and only if it is uniformly continuous if and only if it is bounded. The norm of a linear operator may be computed in several ways, one of which is
\[ \| T \| = \sup \{ \| Tx \| : \| x \| \leq 1 \}. \]

From this it easily seen that whenever the composition \( TS \) is defined, it satisfies \( \| TS \| \cdot \| T \| \| S \| \). One also has \( \| Tx \| \cdot \| T \| \| x \| \) for all \( x \) in \( H \), and in fact

\[ \| T \| = \inf \{ L : \| Tx \| \cdot \| T \| \| x \| \text{ for all } x \text{ in } H \}. \]

The dual space (space of bounded linear functionals) on a Hilbert space may be identified with \( H \) itself, as the following theorem shows.

**Theorem (Riesz Representation Theorem)** Let \( \phi \) be a continuous linear functional on a Hilbert space \( H \). There exists a unique \( y \in H \) such that \( \phi(x) = \langle x, y \rangle \) for all \( x \in H \), and moreover, \( \| y \| = \| \phi \| \).

We leave the proof as an exercise, and suggest that the reader give two proofs, one coordinate free, and the other in terms of a complete orthonormal system (to be defined shortly) for \( H \). The reader should also verify that the mapping which associates to \( \phi \) the unique vector \( y \) of the theorem is an isometric conjugate-linear map of the dual of \( H \) onto \( H \). Thus the dual of \( H \) may be identified (via this map) with \( H \) itself.

A sequence (or net) \( \{x_n\} \) in \( H \) converges in *norm* to \( x \) if \( \| x_n - x \| \) converges to zero. The Cauchy-Schwarz inequality shows that if \( \{x_n\} \) converges in norm to \( x \), then for each \( y \) in \( H \) we also have \( \langle x_n, y \rangle \) convergent to \( \langle x, y \rangle \). In this latter case, we say that \( \{x_n\} \) **converges weakly to** \( x \).

Thus strong convergence in \( H \) implies weak convergence in \( H \). The converse is not true: in \( l^2(\mathbb{N}) \) the sequence \( e_n \) converges weakly to zero (exercise), but is not norm-Cauchy, since \( \| e_n - e_m \| = \sqrt{2} \) unless \( m = n \).

The **weak topology** for \( H \) is the weakest topology (the one with the fewest open sets) on \( H \) that makes each of the functions in the dual space of \( H \) continuous. That is, the weak topology for \( H \) is the smallest topology in which \( \langle x_n, y \rangle \) is convergent to \( \langle x, y \rangle \) for every \( y \) in \( H \). It is not hard to show that if \( H \) is separable, then on any bounded subset \( S \) of \( H \), the weak topology is metrizable. In fact, if \( \{a_i\} \) is a countable dense subset of \( H \) (or even an orthonormal basis of \( H \)), then a sequence or net \( \{x_n\} \) in \( S \) converges weakly to an element \( x \) of \( S \) if and only if \( d(x_n, x) \) converges to zero, where \( d(x, y) \) is given by

\[
d(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^{\|a_i\|}} \langle x - y, a_i \rangle.
\]
A set \( \{x_\gamma \}_\Gamma \) in \( H \) is called an \textbf{orthonormal system} if \( \langle x_\alpha , x_\beta \rangle \) is one when \( \alpha = \beta \), and is zero otherwise. An orthonormal system is said to be \textbf{complete} if its orthogonal complement is zero. The reader should verify that

1. (\textbf{Bessel’s Inequality}) for each \( x \) in \( H \), \( \|x\|^2 \cdot \sum \langle x, x_\gamma \rangle \), with equality if and only if \( \{x_\gamma \}_\Gamma \) is complete,

2. If \( \{x_\gamma \} \) is a complete orthonormal system for \( H \), then for every \( x \) in \( H \),

\[
x = \sum \langle x, x_\gamma \rangle x_\gamma
\]

3. Every Hilbert space has a complete orthonormal system, any two of which have the same cardinality.

The cardinality of an orthonormal system for \( H \) is called the \textbf{Hilbert space dimension of} \( H \). The dimension of \( H \) is countable if and only if \( H \) is separable.

\textbf{Exercises}

1. Prove the Riesz Representation Theorem. Provide a proof using a complete orthonormal system for \( H \), and also give a coordinate free proof.

2. Let \( C \) be a closed convex set in a Hilbert space \( H \), and let \( x_0 \notin C \). Let \( \{x_n\} \subset C \) be a sequence such that the distance from \( x_n \) to \( x \) converges to \( d(x_0 , C) \). Use the parallelogram identity to show that \( \{x_n\} \) is a Cauchy sequence in \( C \) which converges to a point \( x_c \) with \( \|x_c - x_0\| = d(x_0 , C) \). Prove that \( x_c \) is the unique point with this property.

School of Mathematics
Georgia Institute of Technology
5. A Banach space is said to be uniformly convex if $\forall \varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that $\| x \| = \| y \| = 1$ and $\| x - y \| \cdot \varepsilon = \| \frac{x+y}{2} \| < 1 - \delta(\varepsilon)$. 
Prove that the closest point theorem is true in a uniformly convex space.

6. Give an example showing that even if the "existence" portion of the closest point theorem holds in a Banach space, the "uniqueness" part need not hold.

7. Give an example of a closed convex set C in a Banach space such that there is no point $x_0$ with $\| x_0 \| = \inf \{ \| x \| : x \in C \}$.

8. Prove that every Hilbert space has a complete orthonormal system.


10. Establish the result above that says:

    **Theorem** Let M be a closed subspace of a Hilbert space H. Any $x \in H$ may be written uniquely as $x = m + m_\perp$ with $m \in M$ and $m_\perp \in M_\perp$. The point m is the closest point of M to x. The map $P: x \to m$ is linear, idempotent ($P^2 = P$) and of norm 1.

11. Show that if $\{a_n\}$ is a countable dense subset or an orthonormal basis for a separable H, then the equation $d(x,y) = \sum_{i=1}^{\infty} \frac{1}{\| a_i \|} < x - y, a_i >$ defines a metric on each bounded subset S of H, and that convergence in this metric is weak convergence.

### 3. Examples of Linear Operators

You should keep your study of spectral theory (and any other topic we can think of) firmly grounded in examples. We present here several examples to which we shall return frequently. The identity operator on H is denoted throughout by I.

a. Matrices on $\mathbb{C}^n$. Each n x n matrix A defines a linear operator on $\mathbb{C}^n$. The multiplication operators defined in b below are infinite dimensional analogues of diagonal matrices.
b. Multiplication operators on $l^2$. Let $c = \{c_n\}$ be a sequence of complex numbers, and define $M_{c_n} e_n = c_n e_n$. Then $M \sum a_n e_n = \sum c_n a_n e_n$. If $c$ is a bounded sequence, then $M$ is a bounded operator, and if there exists $\delta > 0$ such that $\delta < |c_n|$ for all $n$, then $M$ has a bounded inverse. Each multiplier $c_n$ is an eigenvalue for $M$, and $e_n$ is an eigenvector for $\lambda = c_n$.

c. Multiplication on $L^2$. Let $\phi$ be a measurable function and define $M: L^2 \to L^2$ by $(M\phi)(t) = \phi(t)f(t)$. If $\phi$ is bounded, then $M$ is bounded, and in fact $\|M\| = \|\phi\|$. Under what circumstances is there a function $f$ with $\|f\| = 1$ and $\|Mf\| = \|\phi\|$? Under what conditions on $\phi$ does the operator $M$ have eigenvalues?

d. Shifts on $l^2 = l^2(\mathbb{N})$. We define the left and right shifts in terms of their action on the unit vectors $e_n$. Specifically, $S_R e_n = e_{n+1}$ and $S_L e_n = e_{n-1}$ if $n > 1$, and $S_L e_1 = 0$. These are easily expressed in terms of the action on sequences in $l^2$:

$$S_R(a_1,a_2,a_3,...) = (0,a_1,a_2,a_3,...)$$

$$S_L(a_1,a_2,a_3,...) = (a_2,a_3,a_4,...).$$

One sees immediately that $\|S_L\| = \|S_R\| = 1$. Notice that $S_R$ is 1-1 but not onto (the range has co-dimension 1) while $S_L$ is onto and has a one dimensional kernel. We have that $S_L S_R = I$, but $S_R S_L \neq I$.

e. Shifts on $l^2(\mathbb{Z})$. The right shift on the double-barrelled Hilbert space $l^2(\mathbb{Z})$ is defined just as above, while the left shift satisfies $S_L e_n = e_{n-1}$ for all $n$. In this case $S_L$ and $S_R$ are inverses of each other, and not only do we have $\|S_R\| = \|S_L\| = 1$, but in fact $\|S_R x\| = \|S_L x\| = \|x\|$ for every $x \in l^2(\mathbb{Z})$.

f. Weighted shifts on $l^2$. The weighted shifts are the operators $M_{S_R}$ and $M_{S_L}$ with $M$, $S_R$, and $S_L$ as defined above.
g. Integral operators. Let $K(s,t)$ be a function of two variables, defined say, on the square $[0,1] \times [0,1]$. Assuming $K(s,t)$ satisfies some natural condition, such as continuity or square-integrability, define $T: L^2 [0,1] \rightarrow L^2 [0,1]$ by the formula $(Tf)(t) = \int_0^1 K(s,t) f(s) \, ds$. Check that $T$ is bounded.

h. Differentiation on $L^2 [0,1]$. Define $T$ by $Tf = if'$. There are several things to notice about $T$. First, it is not defined on all of $L^2$. However, we can define it on a dense subspace of $L^2$, such as the domain $D(T) = \{ f : f$ is absolutely continuous, $f' \in L^2$, and $f(0) = f(1) \}$. Second, there is no chance for this operator to be bounded: small functions can have large rates of change. Third, there are many other domains that can be used. Interestingly enough, we shall see that the properties of unbounded operators such as this one depend on the domain as well as the formula of definition.

**Exercise**

1. Compute the norms of the following operators discussed in the examples: 
   - shifts, multiplications, and weighted shifts on $l_2(N)$ and $l_2(Z)$;
   - multiplication on $L^2$.

2. Compute an upper bound on the norm of the integral operator in example g.

4. **Operators and their Adjoints**

   We begin with some observations on equality of operators. To show that $Sx = Tx$ for all $x$, it suffices to show that $S - T = 0$. Since the orthogonal complement of $H$ is $\{ 0 \}$, we have that $S = 0$ if and only if $\langle Sx, y \rangle = 0$ for all $x, y \in H$. For complex Hilbert spaces, even more is true. The next proposition is called the **Polarization Identity**. In the special case when $S = I$, this identity shows how to recapture the inner product on a Hilbert space from its norm.

**Proposition (Polarization Identity)** Let $S$ be an operator on a Hilbert space $H$. Then for any $x, y \in H$,

$$ 4 \langle Sx, y \rangle = \langle S(x + y), (x + y) \rangle - \langle S(x - y), (x - y) \rangle + $$
\[ i \langle S(x + iy), (x + iy) \rangle - i \langle S(x - iy), (x - iy) \rangle. \]

**proof:** Just expand the right hand side.

**Proposition** An operator \( S \) on a complex Hilbert space is 0 if and only if \( \langle Sx, x \rangle = 0 \) for every \( x \in H \).

**proof:** One direction is trivial. The other is an easy consequence of the polarization identity.

We now use the Riesz Representation Theorem to define the **adjoint** of a bounded linear operator. For each \( y \in H \), the formula

\[ x \rightarrow \langle Tx, y \rangle \]

defines a linear functional \( \phi \) on \( H \). By the Cauchy-Schwarz Inequality

\[ |\phi(x)| \leq \| Tx \| \| y \| \leq \| T \| \| y \| \| x \| \]

so \( \phi \) is bounded. Thus there is a unique \( y' \in H \) with

\[ \phi(x) = \langle x, y' \rangle. \]

We define \( T^* \) \( y = y' \). By the uniqueness part of the Riesz Representation Theorem, \( T^* \) is linear. For emphasis we state again the condition satisfied by \( T^* \):

\[ \langle Tx, y \rangle = \langle x, T^* y \rangle \quad \forall x, y \in H. \]

The adjoint \( T^* \) is bounded since

\[ \| T^* \| = \sup \{ \| \langle x, T^* y \rangle \| : \| x \| = \| y \| = 1 \} \]
\[ = \sup \{ \| \langle Tx, y \rangle \| : \| x \| = \| y \| = 1 \} \]
\[ = \| T \|. \]

The method we used in constructing the adjoint is useful in other contexts as well. Here’s one we’ll use later. A **sesquilinear functional** on \( H \) is a function \( \phi : H \times H \rightarrow \mathbb{C} \) which is linear in its first variable and
conjugate linear in the second variable. It is said to be a **bounded sesquilinear functional** if there exists \( M > 0 \) with \( |\phi(x,y)| \leq M \|x\| \|y\| \) for all \( x, y \in H \). The norm of \( \phi \) is defined to be the infimum of the \( M \)'s which work in this inequality. The inner product is an example of a bounded sesquilinear functional. The proof of the next proposition is left to the reader.

**Proposition** If \( \phi \) is a bounded sesquilinear functional on a Hilbert space \( H \), then there is a unique bounded linear operator \( T \) on \( H \) with \( \phi(x,y) = \langle Tx, y \rangle \) for all \( x, y \in H \). In addition, \( \|T\| = \|\phi\| \).

The following Proposition gives a useful relationship between the operators \( T \) and \( T^* \).

**Proposition** Let \( T \in \mathcal{B}(H) \). Then
\[
\text{a. } \mathcal{R}(T) = (\ker(T^*))^⊥ \\
\text{b. } \mathcal{R}(T^*) = (\ker(T))^⊥
\]

**proof:** If \( Tx \in \mathcal{R}(T) \) and \( y \in \ker(T^*) \), then \( \langle Tx, y \rangle = \langle x, T^* y \rangle = 0 \), so \( \mathcal{R}(T) \subset (\ker(T^*))^⊥ \). Since \( (\ker(T^*))^⊥ \) is closed, we have \( \overline{\mathcal{R}(T)} \subset (\ker(T^*))^⊥ \). On the other hand, if \( y \in \mathcal{R}(T)^⊥ \), then \( 0 = \langle Tx, y \rangle = \langle x, T^* y \rangle \) for any \( x \in H \). Thus \( y \in \ker(T^*) \), and \( \mathcal{R}(T)^⊥ \subset \ker(T^*) \). It follows that \( \ker(T^*)^⊥ \subset \mathcal{R}(T)^⊥ = \overline{\mathcal{R}(T)} \). Statement b follows from a.

**Exercises**

1. Prove the Polarization identity.

2. Give an example of a non-zero linear operator on a real Hilbert space for which \( \langle Tx, x \rangle = 0 \) for all \( x \in H \).

3. Compute the adjoints of the operators discussed in the examples of section 3: shifts, multiplications, and weighted shifts on \( l^2(\mathbb{N}) \), and \( l^2(\mathbb{Z}) \); multiplication on \( L^2 \), and the integral operator \( (Tf)(t) = \int K(s,t) f(s) \, ds \).

4. Prove that for any \( T \in \mathcal{B}(H) \), \( \|T^* T\| = \|T\|^2 \).
5. Prove that if \( \phi \) is a bounded sesquilinear functional on \( H \), then there is a unique bounded linear operator \( T \) with \( \phi(x,y) = \langle Tx, y \rangle \) for all \( x, y \in H \). Show that \( \| \phi \| = \| T \| \).

5. Classes of Operators

In this section we introduce several important classes of operators in \( \mathcal{B}(H) \).

We say that \( T \) is **self-adjoint** if \( T = T^* \), and that \( T \) is **normal** if \( TT^* = T^* T \). An operator \( U \) is **unitary** if \( U^* U = UU^* = I \). The reader should verify the following propositions.

**Proposition** An operator \( T \in \mathcal{B}(H) \) is normal if and only if \( \| Tx \| = \| T^* x \| \) for every \( x \in H \).

**Proposition** The following conditions on an operator \( U \in \mathcal{B}(H) \) are equivalent.

- a. \( U \) is unitary
- b. \( \| Ux \| = \| x \| \) for every \( x \in H \) and \( \mathcal{R}(U) = H \).
- c. \( \langle Ux, Uy \rangle = \langle x, y \rangle \) for every \( x, y \in H \) and \( \mathcal{R}(U) = H \).

A **projection** is a bounded linear operator \( P \) which is idempotent, namely, \( P^2 = P \). The range of a projection is always closed. In fact, if \( \{ x_n \} \) is a Cauchy sequence in \( \mathcal{R}(P) \), then \( x_n \to x_0 \) for some \( x_0 \in H \). Since \( P \) is continuous, \( Px_n \to Px_0 \). Now \( x_n \in \mathcal{R}(H) \), so \( x_n = Px_n \), and hence \( x_0 = Px_0 \) is in \( \mathcal{R}(P) \).

Every \( x \in H \) can be written uniquely as \( x = Px + (I - P)x \), with \( Px \in \mathcal{R}(P) \) and \( (I - P)x \in \ker(P) \). In the event that \( \mathcal{R}(P) \perp \ker(P) \), we say that \( P \) is an **orthogonal projection**. A projection is orthogonal if and only if it is self-adjoint.

Let \( B \) denote the unit ball in \( H \), i.e. \( \{ x \in H : \| x \| \cdot 1 \} \). A linear operator \( T \) is called **compact** if the image of the unit ball has compact closure. That is, \( \overline{T(B)} \) is compact. It is not hard to show that \( T \) is compact if and only if the closure of the image of the open unit ball (or any ball, open
or closed, of strictly positive radius) is compact. For Hilbert spaces we have the following useful simplification of the definition.

**Proposition** Let $B$ denote the closed unit ball of a Hilbert space $H$. Let $T \in B(H)$. Then $T$ is compact if and only if $T(B)$ is norm-compact.

**proof:** Suppose $T$ is compact. By the Hahn-Banach theorem the weak closure of the convex set $T(B)$ agrees with its norm closure. Since the common closure $K$ is norm-compact, it is also weakly compact. Now the norm topology contains the weak topology, since both are compact and Hausdorff, they must agree on $K$. The operator $T$ is continuous from the weak topology to the weak topology (exercise). Since $H$ is reflexive, the weak and weak* topologies coincide, so $T$ is weak*-to-weak*-continuous from $B$ into $K$. By Alaoglu’s theorem, $B$ is weak* compact, and hence weakly compact. Thus $T(B)$ is weakly compact, so is weakly closed. But this just says that $T(B) = K$, and hence $T(B)$ is norm compact. The converse is immediate.

We summarize several properties of compact operators in the next Theorem. The proofs are easier in the Hilbert space setting, but all remain true in the more general setting of operators on Banach spaces. In the language we will be using later, this theorem states that the compact operators on $H$ form a closed, two-sided, *-stable ideal in $B(H)$.

**Theorem** Let $T$ be a linear operator on $H$, and let $S \in B(H)$.

a. If $T$ is compact, then $T \in B(H)$.

b. If $\{T_n\}$ is a sequence of compact operators, and $\|T_n - T\| \to 0$, then $T$ is compact.

c. If $R(T)$ is finite dimensional (we say $T$ is a finite rank operator) and $T$ is bounded, then $T$ is compact.

d. A projection is compact if and only if it has finite dimensional range.

e. $T$ is compact if and only if $T^*$ is compact.

f. If $T$ is compact, then both $TS$ and $ST$ are compact.

**proof:** The easiest proofs use the proposition above. We do part e as an example, and leave the rest to the reader. Suppose that $T x_n$ converges in $T(B)$ to $y$. We must show that $y = Tx$ for some $x$ in $B$. Since norm convergence implies weak convergence in $H$, $T x_n$ converges weakly to $y$. On the other hand, by Alaoglu’s Theorem and the fact that $H$ is reflexive, $B$ is weakly compact. Thus we may pass to a convergent subnet (subsequence if $H$ is separable) $x_{n_r}$ of $x_n$. Since $T$ is bounded, $T$ is weak-to-weak continuous...
(exercise), so $T x_n$ converges weakly to $T x$, where $x$ is the weak limit of $x_n$.
Thus $y = T x$ as desired.

**Exercises**

1. Prove that an operator $T \in \mathcal{B}(H)$ is normal if and only if $\| T x \| = \| T^* x \|$ for every $x \in H$.

2. Suppose $U \in \mathcal{B}(H)$ maps $H$ onto itself. Prove that the following statements are equivalent.
   a. $U$ is unitary
   b. $\| U x \| = \| x \|$ for every $x \in H$
   c. $\langle U x, U y \rangle = \langle x, y \rangle$ for every $x, y \in H$.
   Give an example of an operator which is not unitary, yet for which b. and c. hold.

3. Prove that a projection (an idempotent in $\mathcal{B}(H)$) is self-adjoint if and only if it is orthogonal. If $P$ is a self-adjoint projection, describe the operator $I - 2 P$. Give an example of a projection which is not self-adjoint.

4. Let $T \in \mathcal{B}(H)$ be the integral operator $(Tf)(t) = \int \int K(s,t) \, ds$ with $K(s,t) \in L^2([0,1] \times [0,1])$. Show that $T$ is compact.

5. Under what circumstances is a multiplication operator on $l_2(\mathbb{N})$ compact?

6. When is a multiplication operator on $L^2$ compact?

7. Let $T$ be a weighted shift on $l_2(\mathbb{N})$, $T x = \sum a_n \cdot x \cdot e_n \cdot e_{n+1}$, where $a_n \to 0$. Prove that $T$ is compact.

8. Prove the last theorem of this section.

6. **Invertibility and the Spectrum of an Operator**
There's basically only one way for an operator $T$ on a finite dimensional space $X$ to fail to have an inverse. Since

$$\dim(\mathcal{R}(T)) + \dim(\ker(T)) = \dim(X) < \cdot,$$

$T$ fails to be onto precisely when it fails to be 1-1. As the shifts on $l^2$ show, these properties are not linked for operators on infinite dimensional spaces. There are three ways an operator on an infinite dimensional space can fail to have a bounded inverse:

a. It is not 1-1 (Like the left shift)

b. It is not onto (Like the right shift)

c. It's trying hard to have an inverse, but the inverse is not bounded.
(Think of a multiplication operator on $l^2$ given by a sequence $\{a_n\}$ with $a_n \cdot 0 \forall n$, and $\lim_n a_n = 0$.)

**Theorem** Let $T \in \mathcal{B}(H)$. The following are equivalent.

a. $T$ has a bounded inverse.

b. $T^*$ has a bounded inverse.

c. There is a constant $c > 0$ such that $\| Tx \| \cdot c \| x \|$ and $\| T^* x \| \cdot c \| x \|$ for every $x \in H$.

d. $T$ and $T^*$ are both 1-1 and $\mathcal{R}(T)$ is closed.

e. $T$ is 1-1 and onto.

**proof:** It is clear that a is equivalent to b.

If $T$ has a bounded inverse, then $\| x \| = \| T^{-1} Tx \| \cdot \| T^{-1} \| \| Tx \|$, so $\| Tx \| \cdot (\| T^{-1} \| \cdot \| T \|^{-1}) \| x \|$ for all $x$. Since $T^*$ is also invertible, a similar inequality holds for $T^*$, and we may take $c = \min\{ (\| T^{-1} \| \cdot \| T^* \|^{-1}) \}$. Thus a implies c.

An inequality of the type $\| Tx \| \cdot c \| x \|$ clearly shows that $T$ is 1-1. Hence if c. holds, both $T$ and $T^*$ are 1-1. This type of inequality also always implies $\mathcal{R}(T)$ is closed. Indeed, if $y_n = Tx_n$ is a Cauchy sequence in $\mathcal{R}(T)$ converging to $y_0$, then $\{x_n\}$ is even more Cauchy, and thus converges, say to $x_0$, and $Tx_0 = y_0$. Thus c. implies that $\mathcal{R}(T)$ is closed, so c. implies d.

As for d. implies e., we need only show that $\mathcal{R}(T) = H$. By d., $\ker(T^*) = \{0\}$, so since $\mathcal{R}(T)$ is closed, we have $\mathcal{R}(T) = (\ker(T^*))^\perp = H$.

Finally, e. implies a. by the Closed Graph Theorem.
The set of operators with bounded inverses is an open subset of $\mathcal{B}(H)$. The first step towards this result is the so-called Neumann series for the inverse of $(I - T)$, where $T$ is a linear operator of small norm. The second step is to show that an operator close to an invertible operator is invertible.

**Proposition** If $\|T\| < 1$, then $(I - T)$ has a bounded inverse.

**proof:** Since $\|T\| < 1$, the series $\sum_{n=0}^{\infty} T^n$ is convergent in the norm topology of $\mathcal{B}(H)$ to an operator $S$. Furthermore $(I - T)(I + T + T^2 + \cdots + T^k) = (I + T + T^2 + \cdots + T^k)(I - T) = I - Tk+1$. It follows that $S(I - T) = (I - T)S = I$.

**Proposition** If $T$ is an invertible operator and $\|S - T\| < (\|T^{-1}\|)^{-1}$, then $S$ is invertible. Consequently the set of invertible operators is an open subset of $\mathcal{B}(H)$.

**proof:** Write $S = T - (T - S) = T(I - T^{-1}(T - S))$. The hypotheses imply that $\|T^{-1}(T - S)\| < 1$, so $I - T^{-1}(T - S)$ is invertible. It follows that $S$ is invertible.

If $T \in \mathcal{B}(H)$, the **resolvent set** of $T$ is $\rho(T) = \{ \lambda : T - \lambda I$ is invertible$\}$. Since the invertible operators are an open subset of $\mathcal{B}(H)$, the resolvent set of a bounded operator is always an open subset of $\mathbb{C}$. The **spectrum** of $T$ is $\sigma(T) = \mathbb{C} \setminus \rho(T)$, and is always closed.

The complex numbers $\lambda$ in the spectrum of $T$ may be classified according to the way in which $T - \lambda I$ fails to be invertible. The **point spectrum** of $T$, denoted by $\sigma_p(T)$, is the set of all $\lambda$ for which $T - \lambda I$ is not 1-1. That is, the point spectrum consists of the eigenvalues of $T$. The **continuous spectrum** of $T$, $\sigma_c(T)$, is the set of all $\lambda$ for which $T - \lambda I$ is 1-1, and $\mathcal{R}(T - \lambda I)$ is a proper dense subspace of $H$, and the **residual spectrum**, $\sigma_r(T)$ consists of all the points of $\sigma(T)$ left over, namely those $\lambda$ for which $T - \lambda I$ is 1-1 and $\mathcal{R}(T - \lambda I)$ is a proper subspace of $H$. It's not hard to see that for any $T \in \mathcal{B}(H)$, $\sigma(T)$ is the disjoint union of the point, continuous, and residual spectra of $T$. Furthermore, if $H$ is finite dimensional then $\sigma(T) = \sigma_p(T)$. Generally, the way for a point to end up in $\sigma_r(T)$ is to be an
eigenvalue of $T^*$. The **approximate point spectrum**, $\sigma_{ap}(T)$ consists of points which are eigenvalues or "almost" eigenvalues. Specifically, $\sigma_{ap}(T)$ consists of all $\lambda$ such that $\forall \varepsilon > 0$, there is a unit vector $x \in H$ with $\|Tx - \lambda x\| < \varepsilon$.

We'll discuss many properties of the spectrum in the context of Banach algebras. For now we are content with the following observations.

**Proposition** For any $T \in B(H)$, $\sigma_c(T) \subset \sigma_{ap}(T)$.

**proof:** If $\lambda \in \sigma(T)$ but $\lambda \not\in \sigma_{ap}(T)$, then there exists $c > 0$ such that $\|Tx - \lambda x\| \cdot c \|x\|$ for all $x \in H$. But this implies $\mathcal{R}(T - \lambda I)$ is closed, so $\lambda \not\in \sigma_c(T)$.

**Proposition** If $T$ is a bounded normal operator, then $\sigma_r(T) = \phi$.

**proof:** If $\lambda \in \sigma_r(T)$, then $\ker(T^* - \overline{\lambda} I) = \mathcal{R}(T - \lambda I)\perp \cdot \{0\}$. Since $T$ is normal, so is $T - \lambda I$, and thus for any $x$, $\| (T - \lambda I)x \| = \| (T^* - \overline{\lambda} I)x \|$, so $\ker(T - \lambda I) = \ker(T^* - \overline{\lambda} I) \cdot \{0\}$. But this implies $\lambda \in \sigma_p(T)$, a contradiction.

**Examples:**

1) The number 1 is the unique element of the spectrum of the identity $I$, and it lies in the point spectrum of $I$. Every eigenvalue of a matrix lies in the point spectrum of the corresponding operator, and the spectrum of this operator is precisely the set of eigenvalues.

2) If $T$ is multiplication by the function $f(x) = x$ on $L^2[0,1]$, then every element of $[0,1]$ lies in the continuous spectrum of $T$. (In particular, this implies that the elements of $H = L^2[0,1]$ of the form $xh(x)$, where $h$ lies in $H$, form a dense proper subspace of $H$; this can be established by using the Weierstrass approximation theorem.) Moreover the spectrum of this $T$ is precisely $[0,1]$.

3) For each $t$ with $|t| < 1$, the geometric sequence $\{1, t, t^2, \ldots \}$ is an eigenvector for the left shift $S = S_L$. For each such $t$, $(S - tI)^*$ is injective (why?); the range of $(S - tI)^*$ is contained in the orthogonal complement of this eigenvector, so the conjugate of each such $t$ lies in the residual spectrum of $S^* = S_R$.

**Exercises**
1. Let $M$ be a multiplication operator on $l^2$. Compute the point spectrum, continuous spectrum, and residual spectrum of $M$. What is the approximate point spectrum of $M$?

2. Compute the point spectrum, continuous spectrum, residual spectrum, and approximate point spectrum for
   a. the right shift on $l^2(\mathbb{N})$.
   b. the left shift on $l^2(\mathbb{N})$.
   c. the right shift on $l^2(\mathbb{Z})$.
   d. a multiplication operator on $L^2[0,1]$. In particular, when is $\sigma_p \cdot \phi$?

3. What is the spectrum of a projection? Does it matter whether or not the projection is self-adjoint?

4. Work out the assertions given in the list of examples at the end of the section above.

Part II. Banach Algebras

7. Normed algebras and the Gelfand-Mazur Theorem

   An algebra over a field $F$ is a set $A$ equipped with operations of addition, multiplication, and scalar multiplication such that $A$ is a vector space under the operations of addition and scalar multiplication, a ring under addition and multiplication, and an associative law: $\alpha(ab) = (\alpha a)b = a(\alpha b)$ holds whenever $a,b \in A$ and $\alpha \in F$.

   Examples abound. If $S$ is any set, then the space $F_F(S)$ of $F$ valued functions on $S$ with operations defined point-wise is an algebra. So is the space of bounded $F$ valued functions on $S$. The space $C[0,1]$ of complex valued continuous functions on $[0,1]$ and the space of continuous real valued functions $C R [0,1]$ are both algebras. In the preceding examples the multiplication (defined point-wise) is commutative. The space of all square $n \times n$ matrices and the space $B(H)$ (using composition as multiplication) are examples of non-commutative algebras.

   We say a ring is a division ring if it has a multiplicative identity and every non-zero element has a multiplicative inverse. Not surprisingly, if an
algebra is a division ring we call it a division algebra. Examples of division algebras over \( \mathbb{R} \) include \( \mathbb{R} \), \( \mathbb{C} \), and the quaternions. For any field \( F \), the space of all rational functions of one variable is a division algebra over \( F \).

If \( A \) is a division algebra with identity \( I \) and \( a \in A \) is not a scalar multiple of \( I \), then for any \( \lambda \in F \), \( a - \lambda I \cdot 0 \), so \( a - \lambda I \) is invertible. If \( A \) is not a division algebra, for \( a \in A \), there are generally \( \lambda \) for which \( (a - \lambda I) \) is not invertible. The spectrum of \( a \) in \( A \), denoted by \( \sigma_A(a) \) or \( \sigma(a) \) is \( \{ \lambda \in \mathbb{C} : a - \lambda I \) is not invertible\}. We should point out that if \( A \) is a subalgebra of \( B \), and if \( a - \lambda I \) is not invertible in \( B \), then it isn’t invertible in \( A \) either. Thus \( \sigma_B(a) \subset \sigma_A(a) \). The resolvent set of \( a \) is the complement of the spectrum of \( a \).

As remarked earlier, the spectrum of an \( n \times n \) matrix is the set of eigenvalues of the matrix. An example we’ll return to involves the space \( C[0,1] \) (or if you desire greater generality, \( F_{FS} \)). Keep in mind that in \( C[0,1] \) the multiplication is the pointwise multiplication of functions, so \( g^{-1} \) is the pointwise reciprocal of \( g \). If \( g \in C[0,1] \), then \( g^{-1} \) exists \( \iff g \) is never 0. Thus \( f - \lambda I \) is invertible \( \iff f \) never assumes the value \( \lambda \). Thus \( \sigma(f) \) is the range of the function \( f \). In these examples the spectrum is nonempty. This is not generally true, and in particular, the reader should check that in a division algebra either \( \sigma(a) = \emptyset \) or \( a = \lambda I \), in which case \( \sigma(a) = \{ \lambda \} \).

A normed algebra is an algebra which is a normed space and in which \( ||ab|| \cdot ||a|| ||b|| \) for all \( a, b \in A \). If \( A \) contains an identity \( I \), then \( ||I|| = ||II|| \cdot ||I||^2 \), so either \( ||I|| = 0 \) or \( ||I|| \cdot 1 \). Unless explicitly stated to the contrary, we shall assume that \( ||I|| = 1 \). A complete normed algebra is called a Banach algebra.

In a normed algebra, the multiplication operation is jointly continuous.

**Proposition** Let \( A \) be a normed algebra. If \( a_n \to a \) and \( b_n \to b \), then \( a_n b_n \to ab \).

**Proof:** \( ||a_n b_n - ab|| \cdot ||a_n(b_n - b)|| + ||(a_n - a)b|| \cdot ||a_n|| ||b_n - b|| + ||a_n - a|| ||b|| \). These terms both tend to zero since \( \{ ||a_n|| \} \) is bounded and \( a_n \to a \) and \( b_n \to b \).

Much of the discussion of invertibility in Section 6 carries over to the Banach algebra setting.
**Proposition** If $A$ is a Banach algebra with identity $I$, and $\| a \| < 1$, then $I - a$ is invertible and $\| (I - a)^{-1} \| \cdot \frac{1}{1 - \| a \|}$.

proof: The proof of invertibility of $(I - a)$ is exactly the same as in Section 6, and $(I - a)^{-1} = \sum_{n=0}^{\infty} a^n$. The norm estimate follows from the fact that for each $N$, $\| \sum_{n=0}^{N} a^n \| \cdot \sum_{n=0}^{N} \| a^n \| \cdot \frac{1}{1 - \| a \|}$.

Exactly as before it follows that the invertible elements of $A$ form an open set, that $\sigma(a)$ is always a closed subset of $\mathbb{C}$, and furthermore, these ideas show that the spectrum of an element $a \in A$ is bounded.

**Proposition** If $a$ is an element of a Banach algebra and $\lambda \in \sigma(a)$, then $|\lambda| \cdot \| a \|$. Thus $\sigma(a)$ is a compact subset of $\mathbb{C}$.

proof: If $|\lambda| > \| a \|$, then $\| \frac{a}{\lambda} \| < 1$, so $I - \frac{a}{\lambda}$ is invertible. It follows that $\lambda (I - \frac{a}{\lambda}) = \lambda I - a$ is invertible, so that $\lambda \notin \sigma(a)$.

Notice that we really get more out of the proof of this last proposition, since we have a series representation for certain inverses. What we have is

**Proposition** If $a$ is an element of a Banach algebra and $|\lambda| > \| a \|$, then $\lambda I - a$ is invertible, and $(\lambda I - a)^{-1} = \sum_{n=0}^{\infty} \frac{a^n}{\lambda^{n+1}}$.

One can also show that inversion is continuous, as stated in the next proposition. Notice that the invertibility of $a$ is part of the hypothesis, and that from this follows the invertibility of $a_n$ for large $n$. The proof is left to the reader.

**Proposition** If $a$ is invertible and $a_n \rightarrow a$, then $a_n^{-1} \rightarrow a^{-1}$.
We will next show that the spectrum of an element of a Banach algebra $A$ is nonempty. For $a \in A$, the function

$$ r_a(\lambda) = (a - \lambda I)^{-1} $$

is called the **resolvent function of $a$**. It is defined on the resolvent set of $a$, which we have seen is an open subset of $\mathbb{C}$.

The resolvent function is analytic in the sense that for any $\psi \in A^*$, $\psi(r_a(\lambda))$ is analytic. In fact, letting $f(\lambda) = \psi(r_a(\lambda))$, at any $\lambda_0$ in the resolvent set of $a$,

$$ \frac{f(\lambda) - f(\lambda_0)}{\lambda - \lambda_0} = \psi\left(\frac{(a - \lambda I)^{-1} - (a - \lambda_0 I)^{-1}}{\lambda - \lambda_0}\right) $$

$$ = \psi\left((a - \lambda I)^{-1}\frac{(a - \lambda_0 I) - (a - \lambda I)}{\lambda - \lambda_0}(a - \lambda_0 I)^{-1}\right) $$

$$ = \psi\left((a - \lambda I)^{-1}(a - \lambda_0 I)^{-1}\right) $$

Sending $\lambda \to \lambda_0$, we obtain (using the continuity of inversion and of $\psi$) that

$$ \frac{df}{d\lambda} = \psi((a - \lambda_0 I)^2) $$

so $r_a(\lambda)$ is an analytic function on the resolvent set of $a$.

Furthermore,

$$ r_a(\lambda) = (a - \lambda I)^{-1} = \frac{1}{\lambda} \left(\frac{a}{\lambda} - I\right)^{-1} $$

Now as $|\lambda| \to \bullet$, $\frac{1}{\lambda} \to 0$, so $(\frac{a}{\lambda} - I)^{-1} \to -I$ by continuity of inversion. It then follows that $r_a(\lambda) \to 0$ as $\lambda \to \bullet$.

Now suppose $\sigma(a) = \emptyset$. Then for any $\psi \in A^*$, $\psi(r_a(\lambda))$ is an entire function which tends to zero as $|\lambda| \to \bullet$. Thus $\psi(r_a(\lambda))$ is bounded, so for
any $\psi \in A^*$, $\psi(r_a(\lambda))$ is a bounded entire function, and is thus constant by Liouville’s Theorem. Since $r_a(\lambda) \to 0$ as $\lambda \to \cdot$, $\psi(r_a(\lambda)) = 0$ for all $\psi$, so $r_a(\lambda) \equiv 0$. But 0 isn’t the inverse of anything, so we have a contradiction. We have thus proved the following theorem.

**Theorem** Let $a$ be an element in a Banach algebra with identity. The spectrum of $a$ is a compact nonempty subset of $C$.

An **isomorphism** between $A$ and $B$ is generally a 1-1 and onto map which preserves all of the relevant structures. Thus a ring isomorphism preserves multiplication and addition, an isomorphism of Banach spaces preserves the vector space and metric structures, and an isomorphism of circuses maps green-eyed lions to green-eyed lions. The next theorem deals with isomorphisms of division algebras.

**Theorem (Gelfand-Mazur Theorem)** If $A$ is a normed algebra over $C$ which is a division ring, then $A$ is isometrically isomorphic to $C$.

**proof:** Let $a \in A$. By exercise 6 below, $A$ has a completion which is a Banach algebra. The spectrum of $a$ in this completion is not empty. Thus the spectrum $\sigma(a)$ of $a$ in $A$ cannot be empty, since $A$ is a subalgebra of its completion. Thus there exists $\lambda_a \in \sigma(a)$. Since $a - \lambda_a I$ is not invertible and since $A$ is a division ring, $a - \lambda_a I = 0$, so $a = \lambda_a I$. It is easy to check that the map $a \to \lambda_a$ is an isomorphism of algebras, and since $\| I \| = 1$, we have $\| a \| = \| \lambda_a I \| = \| \lambda_a \|$ so the map is isometric.

**Exercises**

1. Prove that if $\{a_n\}$ is a sequence in a Banach algebra, $a_n \to a$, and $a$ is invertible, then $a_n^{-1} \to a^{-1}$.

2. Let $A$ be a Banach algebra without an identity, and define $A_1 = A \times C$, with norm $\| (x, \alpha) \| = \| x \| + |\alpha|$ and operations defined by
   a. $(x, \alpha) + (y, \beta) = (x + y, \alpha + \beta)$ $\forall x, y \in A$ and $\alpha, \beta \in C$.
   b. $\beta (x, \alpha) = (\beta x, \alpha \beta)$ $\forall x \in A, \alpha, \beta \in C$.
   c. $(x, \alpha) (y, \beta) = (xy + \alpha y + \beta x, \alpha \beta)$ $\forall x, y \in A$ and $\alpha, \beta \in C$.

   Prove that $A$ is a Banach algebra with identity $(0,1)$, and that $A$ is isometric to a subalgebra of $A_1$. What happens if this process for adding
8. Ideals in Banach Algebras

An (two-sided) **ideal** in an algebra $A$ over a field $F$ is a vector subspace $I$ of $A$ which has the property that whenever $a \in A$ and $b \in I$, $ab \in I$ and $ba \in I$. An ideal $I$ is **proper** if $I \neq A$, and **maximal** if whenever $J$ is a proper ideal with $I \subset J \subset A, I = J$. It is clear that if an ideal $I$ contains an invertible element of $A$, then $I = A$.

Using Zorn's lemma, we can show that there are generally lots of maximal ideals.

**Proposition** If $I$ is an ideal in an algebra $A$ with identity $I$, there is a maximal ideal $M$ containing $I$. 

proof: Let \( J \) denote the collection of all proper ideals which contain \( I \), ordered by inclusion. If \( \mathcal{C} \) is a totally ordered subset of \( \mathcal{C} \), then \( \mathcal{C}_0 \) = \( \mathcal{C} : C \in \mathcal{C} \) is an ideal. Since it does not contain \( I \), it is proper, and is an upper bound for \( \mathcal{C} \). By Zorn's lemma, \( J \) contains a maximal element.

**Proposition** If \( A \) is a Banach algebra with identity \( I \), every maximal ideal in \( A \) is closed.

proof: It is easy to check that if \( I \) is an ideal in \( A \), then \( \overline{I} \) is also an ideal. Since the set of invertible elements is open in \( A \), if \( I \) is a proper ideal in \( A \), \( \overline{I} \) is also a proper ideal. Thus if \( I \) is maximal we must have \( I = \overline{I} \).

Whenever \( I \) is an ideal in an algebra \( A \), the quotient space \( A/I \) is an algebra. The multiplication of cosets, \((a + I)(b + I) = ab + I\), is well defined since \( I \) is an ideal, and the coset \( I + I \) is an identity for \( A/I \). If \( A \) is a normed algebra, we may define a norm on the quotient space by

\[
\|a + I\| = \inf \{ \|a + j\| : j \in I \}.
\]

It is easy to check that the map \( a \mapsto a + I \) is a norm decreasing homomorphism of algebras. If \( A \) is complete and \( I \) is closed, then \( A/I \) is also a Banach algebra.

**Theorem** Let \( A \) be a commutative normed algebra with identity over \( \mathbb{C} \), and let \( M \) be a maximal ideal in \( A \). Then \( A/M \) is isomorphic to \( \mathbb{C} \).

proof: We know that \( A/M \) is a field, and hence a division algebra. Since \( A \) is normed, so is \( A/M \), so by the Gelfand-Mazur Theorem, \( A/M \) is isomorphic to \( \mathbb{C} \).

A linear functional \( \phi \) on a Banach algebra \( A \) is a **multiplicative linear functional** if \( \phi(ab) = \phi(a)\phi(b) \) for all \( a, b \in A \). It is clear that if \( A \) has an identity \( I \), then either \( \phi(I) = 1 \) or \( \phi(I) = 0 \). In the latter case, \( \phi \equiv 0 \). We now establish a correspondence between the maximal ideals of \( A \) and the multiplicative linear functionals on \( A \).

**Proposition** Let \( M \) be a maximal ideal in a Banach algebra with identity. Then for each \( a \in A \), there is a unique complex number \( \lambda_a \) such that \( a - \lambda_a \ I \in M \).
proof: Since $A/M$ is isomorphic to $C$, choose an algebra isomorphism $\tau$ (how many are there?) from $A/M$ onto $C$. Now $\tau$ is a non-zero multiplicative linear functional, so $\tau(I + M) = 1$. Letting $\lambda_a = \tau(a + M)$ we see that $\tau(\lambda_a I + M) = \lambda_a \tau(I + M) = \lambda_a$. Since $\tau$ is 1-1, $a + M = \lambda_a I + M$.

But this means that $a - \lambda_a I \in M$. Thus we’ve shown the existence of the $\lambda_a$. Uniqueness is left to the reader.

**Proposition** Let $M$ be a maximal ideal in a Banach algebra with identity. The mapping $\phi_M$ defined by $\phi_M(a) = \lambda_a$, with $\lambda_a$ as above, is a non-zero multiplicative linear functional on $A$.

proof: The maps $a \rightarrow a + M$ and $a + M \rightarrow \lambda_a$ are both multiplicative and linear, and $\phi_M$ is their composition. Furthermore, $\phi_M(I) = 1$, so $\phi_M$ is non-zero.

Thus each maximal ideal corresponds to a multiplicative linear functional. The converse is also true.

**Proposition** If $\psi$ is a non-zero multiplicative functional on a Banach algebra $A$ with identity, then $\ker(\psi)$ is a maximal ideal in $A$.

proof: Since $\psi$ is linear, $\ker(\psi)$ is a vector subspace of $A$, and since $\psi$ is multiplicative, $a \in A, b \in \ker(\psi)$ implies that $ab$ and $ba$ belong to $\ker(\psi)$. Thus $\ker(\psi)$ is an ideal. As remarked earlier, since $\psi$ is non-zero, $\psi(I) = 1$.

Now suppose $J$ is an ideal in $A$ and $\ker(\psi) \subset J$. Assume that $\ker(\psi) \cdot J$.

We must show $J = A$. To this end, choose $x \in J \setminus \ker(\psi)$. We have then that $\psi(x) \cdot 0$, and $\psi(x - \psi(x) I) = \psi(x) - \psi(x) 1 = 0$. But this implies $x - \psi(x) I \in \ker(\psi) \subset J$, and since $x \in J$, we have that $-\psi(x) I$, and hence $I$, belongs to $J$. Therefore $J = A$.

Summarizing, we have shown that the maps $M \rightarrow \phi_M$ and $\phi \rightarrow \ker(\phi)$ are inverses of each other, and the space of maximal ideals in $A$ is in 1-1 correspondence with the space of non-zero multiplicative linear functionals on $A$. The non-zero multiplicative linear functionals on $A$ are also called characters, and the set of all of them is the character space. We shall denote it by $\chi(A)$. As a corollary, $\phi(a) = 0$ for all $\phi \in \chi(A)$ if and only if $a$ lies in the intersection of the maximal ideals of $A$. This intersection is an ideal in $A$, and is called the radical of $A$. If the radical is $\{0\}$, we say that $A$ is semisimple.

**Exercises**
1. Let $X$ be a compact Hausdorff space, and let $C(X)$ denote the algebra of all continuous complex-valued functions on $X$. Let $J$ be a proper ideal in $C(X)$.

a. Show that there exists an element $x_J \in X$ such that $f(x_J) = 0$ for all $f \in J$. (Suggestion: If not, then for each $x$, there is a function $g_x \in J$ with $g_x(x) > 0$. By compactness, there exists $g \in J$ with $g(x) > 0$ for all $x \in X$.)

b. Show that if $J$ is maximal, then $J = \{ f \in C(X) : f(x_J) = 0 \}$.

c. Show that any multiplicative linear functional $\phi$ on $C(X)$ is of the form $e_x$, where $e_x(f) = f(x)$. Show that the function $x \mapsto e_x$ is a homeomorphism from $X$ onto the character space (with the weak* topology) of $C(X)$.

2. Show that the algebra of $n \times n$ matrices with complex entries has no nontrivial ideals.

9. An Example - Convolution on $l_1(Z)$

We digress briefly from our analysis of general Banach algebras with an important example. We let $l_1 = l_1(Z) = \{ f : Z \to C : \sum_{n=-\infty}^{\infty} |f(n)| < \bullet \}$, with addition defined pointwise, and normed by $||f|| = \sum_{n=-\infty}^{\infty} |f(n)|$. We will denote by $\{e_n\}$ the usual unit vectors ($e_n(j) = \delta_{nj}$), and will generally write the elements of $l_1$ as $f = \sum_{n=-\infty}^{\infty} f(n) e_n$.

We seek a multiplication $\ast$ on $l_1$ under which $l_1$ is a Banach algebra, and which reflects the structure of the underlying space $Z$ of integers. To reflect the (additive) group structure of $Z$, we should require
that \( e_{n*} e_m = e_{n+m} \). It then follows (you check the convergence of the various series) that

\[
\begin{align*}
\sum_{n=-\cdot} f(n) e_n & \ast \left( \sum_{m=-\cdot} g(m) e_m \right) \\
\sum_{n=-\cdot} f(n) g(m) e_{n+m} & = \sum_{j=-\cdot} \left( \sum_{m=-\cdot} f(n) g(j-n) \right) e_j \\
\sum_{k=-\cdot} f(k-m) g(m) & = \sum_{k=-\cdot} e_k.
\end{align*}
\]

This multiplication looks a lot like multiplication of long polynomials, but it’s called \textbf{convolution}. You should check that whenever \( f, g \in l_1 \), we have

\[
\sum_{m=-\cdot} |f(k-m) g(m)| \quad \text{converges, so that the above calculations make sense,}
\]

\( l_1(\mathbb{Z}) \) is then a commutative Banach algebra, with identity \( e_0 \).

Much of the above can be done for a general group \( G \) in place of the integers. One must be careful with the series (or integrals) which arise, but the group structure of \( G \) is reflected in \( l_1(G) \) (or \( L^1(G) \)), and the algebras may be used to study the groups.

We close this section with a discussion of the character space of \( l_1(\mathbb{Z}) \).

\textbf{Theorem} For each \( z \in \mathbb{C} \) with \( |z| = 1 \), the functional \( \phi \) defined on \( l_1 \)

\[
\phi_z(f) = \sum_{n=-\cdot} f(n) z^n
\]

is a non-zero multiplicative linear functional.

Conversely, if \( \phi \in \chi(l_1(\mathbb{Z})) \), there is a unique \( z \) with \( |z| = 1 \) and \( \phi = \phi_z \).
proof: It's easy to see that for each \( z \) of modulus 1, \( \phi_z \) is linear, nonzero, and multiplicative. As for the converse, let \( \phi \) be a non-zero multiplicative linear functional on \( l_1 \). We have \( \phi(e_0) = 1 \), and for \( f \in l_1 \), \( \phi(f - \phi(f)e_0) = 0 \), so \( f - \phi(f)e_0 \) is not invertible in \( l_1 \). Thus \( \phi(f) \in \sigma(f) \), so \( |\phi(f)| \leq \|f\| \). Hence \( \phi \) is bounded (in fact \( \|\phi\| = 1 \)). Hence for any \( n \), \( |\phi(e_n)| \leq 1 \). On the other hand, \( 1 = |\phi(e_0)| = |\phi(e_n^* e_{-n})| = |\phi(e_n)| \cdot |\phi(e_{-n})| \), so \( |\phi(e_n)| \leq 1 \) for all \( n \). The upshot is that \( |\phi(e_n)| = 1 \) for all \( n \). Now let \( z = \phi(e_1) \). Since \( e_n = e_1^n \), it follows that \( \phi(e_n) = z^n \) for all \( n \), and thus that \( \phi = \phi_z \).

The reader may wish to prove the following result, known as Wiener's lemma.

**Proposition** Suppose \( f(e^{i\theta}) = \sum_n a_n e^{i\theta} \) with \( \sum_n |a_n| < \infty \). Suppose that \( f(e^{i\theta}) \neq 0 \) for all \( e^{i\theta} \). Then there exists a sequence \( \{b_n\} \) with \( \frac{1}{f(e^{i\theta})} = \sum_n b_n e^{i\theta} \) and \( \sum_n |b_n| < \infty \).

**Exercise**

1. Prove Wiener's lemma.

**10. The Gelfand Transform**

We begin with a useful representation of a Banach space \( X \) as a subspace of a space \( C(K) \) of continuous functions. Let \( K \) denote the closed unit ball \( \{ x^* \in X^* : \|x^*\| \leq 1 \} \) in \( X^* \) endowed with its weak* topology. Then \( K \) is a Hausdorff (even metrizable if \( X \) is separable) space and is compact by Alaoglu's theorem. We have a natural map \( \Phi \) from \( X \) into the space \( C(K) \) of continuous functions on \( K \) given by \( \Phi(x)(x^*) = x^*(x) \). Since \( K \) is given the weak* topology, \( \Phi(x) \) is a continuous function on \( K \), and \( \Phi \) is easily seen to be an isometry from \( X \) onto a closed subspace of \( C(K) \). (In the event that \( X \) is separable, one can use this to show that \( X \) is isometrically isomorphic to a subspace of \( C[0,1] \). The next step in the argument is to show there's a continuous function from the Cantor set onto \( K \).)

We might ask if such an embedding is possible for a Banach algebra \( A \). We certainly can't find an isometry of \( A \) into a \( C(K) \) algebra unless \( A \) is
commutative, and even in this case $A$ is not always isometric to a subalgebra of a $C(K)$. In this section we develop the Gelfand theory of commutative Banach algebras, and obtain such an isometry for a class of algebras.

Imitating the Banach space situation, we look for a map from $A$ into the continuous functions on the unit ball $K$ of $A^*$. We define the function $\Gamma$ on $A$ by $\Gamma(a)(a^*) = a^*(a)$. This map is linear, but for it to preserve the multiplication of $A$ we must have $a^*(ab) = \Gamma(ab)(a^*) = \Gamma(a)(a^*)\Gamma(b)(a^*) = a^*(a)a^*(b)$. That is, $a^*$ must be a multiplicative linear functional on $A$. Thus the appropriate map to consider is not from $A$ into $C(K)$, but from $A$ into $C(\chi(A))$, the continuous functions on the character space of $A$. Now $K$ is compact and Hausdorff in the weak* topology, and to see that $\chi(A)$ is compact it suffices to show that it is weak* closed in $K$. If $\{\phi_\gamma\}$ is a net in $\chi(A)$ which converges weak* to $\phi$, we have that $\phi(ab) = \lim\phi_\gamma(ab) = \lim(\phi_\gamma(a)\phi_\gamma(b)) = \phi(a)\phi(b)$, so that $\phi$ is multiplicative. Thus $\chi(A)$, equipped with the weak* topology is compact. It is certainly Hausdorff.

We are now in a position to define the Gelfand transform. If $A$ is a commutative Banach algebra with identity, the Gelfand transform of $A$ is the map $\Gamma: A \to C(\chi(A))$ defined by $\Gamma(a)(\phi) = \phi(a)$. Frequently $\Gamma(a)$ is denoted by $^\wedge a$. $\Gamma$ is certainly linear, and we have shown above that it is multiplicative. In addition, $\|\Gamma(a)\| = \sup \{ |^\wedge a(\phi) | : \phi \in \chi(A) \} = \sup \{ |\phi(a)| : \phi \in \chi(A) \} \cdot \|a\|$, so $\Gamma$ is continuous, with norm no larger than 1.

In the last section we identified to character space of $l_1(Z)$ with the unit circle in the complex plane. In fact, the map establishing this correspondence was given by $z \to \phi_z$ where $\phi_z(e^1) = z$. Another way to see this correspondence is via $\Gamma(e^1) = \hat{\phi_z}$, since $\Gamma(e^1)(\phi_z) = \hat{\phi_z}(e^1) = z$. We also obtain additional information. Since $\Gamma(e^1)$ is continuous from a compact space, $\chi(l_1)$, onto a Hausdorff space (the unit circle), we have established that $\chi(l_1)$ is homeomorphic to the unit circle. You should check that for $f \in l_1$, $(\Gamma(f)(\phi_z)) = \sum f(n)z^n$.

We conclude this section with a few observations concerning the Gelfand transform $\Gamma$ and invertibility. You should not be mislead by their apparent simplicity. They are important.

The ideal generated by an element $a \in A$ is the smallest ideal of $A$ which contains $a$. This can also be expressed as the intersection of all ideals containing $a$, and if $A$ is commutative, is $Aa = \{ ba : b \in A \}$. If $A$ is
commutative and has an identity $I$, then $a \in A$ is invertible $\iff$ the ideal generated by $a$ is itself $A$. Why? If $a$ is invertible, then $I = a^{-1} \in A$, so $Aa = A$. On the other hand, if $Aa = A$, then there exists $b \in A$ with $ab = ba = I$. It follows from this that $a$ is invertible $\iff a$ lies in no maximal ideal of $A$.

**Theorem** Let $A$ be a commutative Banach algebra with identity $I$, and let $a \in A$. The following statements are equivalent
a. $a$ is invertible
b. $\Gamma(a)$ is an invertible function in $C(\chi(A))$
c. For all $\phi \in \chi(A)$, $\phi(a) \cdot 0$.
Consequently $\lambda \in \sigma(a)$ if and only if there exists a character $\phi$ with $\phi(a) = \lambda$. Additionally, $\phi(a) = \lambda \iff \Gamma(a)(\phi) = \lambda$, so the range of $\Gamma(a)$ is $\sigma(a)$.

proof: We know that $a$ is invertible $\iff a$ lies in no maximal ideal of $A \iff \phi(a) \cdot 0$ for all $\phi \in \chi(A)$. Since $\Gamma(a)$ is simply a continuous function, $\Gamma(a)$ is invertible $\iff \Gamma(a)(\phi) = \phi(a) \cdot 0$ for all $\phi \in \chi(A)$. Thus a, b, and c are equivalent. Furthermore, from this argument, $a - \lambda I$ is not invertible $\iff$ there exists $\phi \in \chi(A)$ with $\phi(a - \lambda I) = \Gamma(a - \lambda I)(\phi) = 0$. That is $\Gamma(a)(\phi) = \lambda$, so $\lambda$ is in the range of $\Gamma(a)$.

**Theorem** Let $A$ be a commutative Banach algebra with identity $I$. Then $\ker(\Gamma)$ is the radical of $A$. Consequently, $\Gamma$ is 1-1 if and only if $A$ is semisimple.

proof: An immediate consequence of preceding results.

Thus whenever $A$ is commutative and semisimple, $\Gamma$ provides us with a 1-1 homomorphism from $A$ into $C(\chi(A))$. In the next section we will develop a condition that guarantees that $\Gamma$ has dense range in $C(\chi(A))$, and in an important case see that $\Gamma$ is onto as well.

### 11. Involutions and $*$-algebras

An **involution** on a normed algebra $A$ is a function $*$ from $A$ to $A$ which satisfies
a. $a^{**} = a \forall a \in A$,
b. $(a + \lambda b)^* = a^* + \overline{\lambda} \ b^* \ \forall a, b \in A, \lambda \in C$,
c. $(ab)^* = b^*a^* \ \forall a,b \in A$,
d. $|| a^* || = || a ||$.
The easiest example is complex conjugation on $\mathbb{C}$, and as we'll see below, we should always think of $*$ as being somewhat like complex conjugation. Other examples are $f^*(t) = \overline{f(t)}$ on an algebra of continuous functions, $B^* = \text{conjugate transpose of } B$ on the algebra of $n \times n$ matrices, the adjoint operation on $\mathcal{B}(\mathcal{H})$, and $f^* = \sum \overline{f(-n)} e_n$ on $l_1(\mathbb{Z})$.

A normed algebra with an involution is called a **normed $*$-algebra**, and if it is complete we call it a **Banach $*$-algebra**. We usually refer to $a^*$ as "a adjoint". An element $a$ which satisfies $a = a^*$ is called **self-adjoint**, and if $aa^* = a^*a$, $a$ is referred to as a **normal** element of $A$. You should observe that $I$ is self-adjoint, and compute the self-adjoint elements of $l_1(\mathbb{Z})$.

When $L$ is an algebra homomorphism or an isomorphism between $*$-algebras, and $L(a^*) = L(a)^*$, we will call $L$ a $*$-homomorphism or a $*$-isomorphism.

**Proposition** Let $a$ be an element of a normed $*$-algebra $A$ with identity $I$. Then $\lambda \in \sigma(a) \iff \overline{\lambda} \in \sigma(a^*)$.

**proof:** Let $\lambda \in \mathbb{C}$. Then $(a - \lambda I) b = b (a - \lambda I) = I \iff b^* (a - \lambda I)^* = (a - \lambda I)^* b^* = I \iff b^* (a^* - \overline{\lambda} I) = (a^* - \overline{\lambda} I) b^* = I$. Thus $\lambda$ is in the resolvent set of $a$ $\iff \overline{\lambda}$ is in the resolvent set of $a^*$.

**Proposition** Any element $c$ in a $*$-algebra $A$ may be written uniquely as $c = a + ib$, with $a$ and $b$ self-adjoint.

**proof:** Existence and uniqueness follow by solving the system of equations

\[
\begin{align*}
\begin{align*}
a + ib &= c \\
\overline{a} - ib &= \overline{c}.
\end{align*}
\end{align*}
\]

You'll find that $a = \frac{c + \overline{c}}{2}$ and $b = \frac{c - \overline{c}}{2i}$. These ought to look familiar.

**Proposition** If $A$ is a commutative Banach $*$-algebra, and $\psi$ is a continuous linear functional (not necessarily multiplicative) on $A$, then $\psi$ is real on each self-adjoint element of $A$ $\iff \psi(a^*) = \overline{\psi(a)}$ for every $a \in A$. 

32  School of Mathematics  
Georgia Institute of Technology
proof: The $\Leftarrow$ direction is clear. As for the other implication, assume $\psi(a)$ is real whenever $a$ is self-adjoint, write $a = a_1 + i a_2$, with $a_1$ and $a_2$ self-adjoint. Then

$$\psi(a^*) = \psi(a_1 - i a_2) = \psi(a_1) - i \psi(a_2)$$

$$= \psi(a_1) + i \psi(a_2) = \psi(a_1 + i a_2) = \psi(a).$$

**Theorem** Suppose $A$ is a commutative Banach *-algebra with identity $I$, and suppose that every element of $\chi(A)$ is real on the self-adjoint elements of $A$. Then $\Gamma$ is a norm decreasing *-homomorphism from $A$ to a dense *-subalgebra of $C(\chi(A))$.

proof: We already know that $\Gamma$ is a norm decreasing algebra homomorphism. Since each $\phi \in \chi(A)$ is real on self-adjoint elements, $\phi(a^*) = \overline{\phi(a)}$, which means that $\Gamma(a^*)(\phi) = \overline{\Gamma(a)(\phi)}$ for all $\phi \in \chi(A)$. Thus $\Gamma$ is a *-homomorphism. Then $\Gamma(A)$ is a subalgebra of $C(\chi(A))$ which contains the function $1$, separates the points of $\chi(A)$, and contains $\overline{f}$ whenever $f \in \Gamma(A)$. By the Stone-Weierstrass Theorem, $\Gamma(A) = C(\chi(A))$.

If we could show that $\Gamma$ were an isometry, we would then have that $\Gamma(A)$ is closed, and the above theorem would imply that $\Gamma(A) = C(\chi(A))$. Our next goal is to find a condition which ensures that $\Gamma$ preserves norms.

**Exercise**

1. What are the self-adjoint elements of $l_1(\mathbb{Z})$?

**12. The spectral radius formula and the Gelfand-Naimark theorem.**

As advertised, we plan to determine conditions which guarantee that $\Gamma$ preserves norms and that whenever $a = a^*$ we have $\phi(a) \in \mathbb{R}$ for every $\phi \in \chi(A)$. Notice that in $C(\chi(A))$, $\| f \|_1 = \| f \|^2$. If $\Gamma$ preserves norms, then

$$\| a^* a \| = \| \Gamma(a^* a) \| = \| \overline{\Gamma(a)} \Gamma(a) \| = \| \Gamma(a) \|^2 = \| a \|^2.$$

Thus a
necessary condition that Γ preserve norms is that for every \( a \in A, \| a^*a \| = \| a \|^2 \). A Banach *-algebra for which \( \| a^*a \| = \| a \|^2 \) \( \forall a \in A \) is called a \textbf{C*-algebra}, and we will refer to \( \| a^*a \| = \| a \|^2 \) as the \textit{C*-identity}. It should be clear to you that the \textit{C*-identity} implies that \( \| a \| = \| a^* \| \).

We turn now to the condition that \( \phi(a) \) be real for self-adjoint \( a \). It turns out that this always holds in a \textit{C*-algebra}. Since \( \lambda \in \sigma(a) \iff \bar{\lambda} \in \sigma(a^*) \). Since \( a = a^* \), we have \( \lambda \in \sigma(a) \iff \bar{\lambda} \in \sigma(a) \). Letting \( \lambda = x + iy \in \sigma(a) \), we set out to prove that \( y = 0 \), and by these introductory remarks we may assume \( y \cdot 0 \). Since

\[
(a + niI) - (x + (y + n)iI) = a - (x + iy)I,
\]

we see that \( x + (y + n)i \in \sigma(a + niI) \). It then follows that \( x - (y + n)i \in \sigma(a - niI) \). But this implies that

\[
| x + (y + n)i | \quad \cdot \quad \| a + niI \|, \quad \text{and}
\|

| x - (y + n)i | \quad \cdot \quad \| a - niI \| = \| \phi(a + niI)^* \|
\leq \| a + niI \| \quad \text{(by the \textit{C*-identity}).}
\]

Multiplying these inequalities yields

\[
x^2 + y^2 + 2yn + n^2 \quad \cdot \quad \| a + niI \|^2
\]

\[= \| (a + niI)^* (a + niI) \| \quad \text{(C*-identity)}
\]

\[= \| (a - niI) (a + niI) \|
\]

\[\cdot \quad \| a^2 \| + n^2.
\]

Thus \( 0 \cdot 2yn \cdot \| a^2 \| - (x^2 + y^2) \) for all \( n \), and this is clearly impossible.
We return to the problem of showing \( \Gamma \) to be an isometry, which is the case for commutative C*-algebras. Now if \( \Gamma \) is to be an isometry, we must have \( \forall \ a \in A, \)

\[
\| a \| = \| \Gamma(a) \|_\bullet = \sup\{ | \phi(a) | : \phi \in \chi(A) \} \\
= \sup\{| \lambda | : \lambda \in \sigma(a)\}.
\]

Thus our next goal is to show that in any commutative C*-algebra, \( \| a \| = \sup\{| \lambda | : \lambda \in \sigma(a)\} \). This latter number, \( \sup\{| \lambda | : \lambda \in \sigma(a)\} \), is called the **spectral radius** of \( a \), and is the radius of the smallest closed disk centered at 0 which contains \( \sigma(a) \). We shall denote it by \( r(a) \).

**Theorem (Spectral Radius Formula)** If \( A \) is a Banach algebra with identity, and \( a \in A \), then

\[
r(a) = \lim ( \| a^n \|^{1/n} ).
\]

**proof:** We have most of the ingredients already. We have

\[
\lambda^n I - a^n = (\lambda I - a)(\lambda^{n-1} + \ldots + a^{n-1})
\]

so that \( \lambda \in \sigma(a) \) implies \( \lambda^n \in \sigma(a^n) \) (Check this out). Thus \( \lambda \in \sigma(a) \Rightarrow | \lambda | \cdot \| a^n \| \) for all \( n \). That is, | \lambda | \cdot \| a^n \|^{1/n} \) for all \( n \). But this implies that \( r(a) \cdot \lim \| a^n \|^{1/n} \).

Next, the resolvent function \( r_a(\lambda) = (\lambda I - a)^{-1} \) is analytic on the complement of \( \sigma(a) \), which we choose to express by saying \( r_a\left(\frac{1}{z}\right) \) is analytic on \( \{ z : \frac{1}{z} \in \text{resolvent set of } a \} \). Now we have a series representation for \( r_a(\lambda) \), from which we see that

\[
r_a\left(\frac{1}{z}\right) = \sum_{n=0}^{\bullet} a^n z^{n+1}.
\]
We know from the region of analyticity of $r_a \left( \frac{1}{z} \right)$ that this series converges to $r_a \left( \frac{1}{z} \right)$ for $\left| \frac{1}{z} \right| > r(a)$. That is, whenever $|z| < \frac{1}{r(a)}$, the series above converges. It follows that the radius of convergence of this series, $R$, must satisfy $\frac{1}{r(a)} \cdot R$. Now the theory of complex variables gives us another way to calculate $R$, namely, $R = \left( \lim ||a^n||^{1/n} \right)^{-1}$. It follows that $r(a) \cdot \lim ||a^n||^{1/n}$.

Since we’ve established that $r(a) \cdot \lim ||a^n||^{1/n}$ and that $r(a) \cdot \lim ||a^n||^{1/n}$, we have that $r(a) = \lim ||a^n||^{1/n}$.

The reader should note that the first part of the argument now shows that this limit is actually the inf.

We now can prove the (little) Gelfand-Naimark theorem.

**Theorem (little Gelfand-Naimark theorem)** Let $A$ be a commutative $\mathcal{C}^*$-algebra with identity. Then the Gelfand transform $\Gamma$ is an isometric $\ast$-isomorphism from $A$ onto $C(\chi(A))$.

**proof:** Since $A$ is a $\mathcal{C}^*$-algebra, every character of $A$ is real on the self-adjoint elements of $A$, so the last result of the section 11 implies that $\Gamma$ is a norm decreasing (non-increasing, actually) $\ast$-homomorphism from $A$ onto a dense subalgebra of $C(\chi(A))$.

If $a$ is self adjoint, then $||a^2|| = ||a^*a|| = ||a||^2$. Continuing, we get $||a^{2n}|| = ||a||^{2n}$ whenever $a$ is self-adjoint, so that for self-adjoint $a$, the norm $||a||$ and spectral radius $r(a)$ are equal. Thus

$$||\Gamma(a)|| = \sup\{ |\Gamma(a)(\phi)| \} = r(a)$$

$$= \lim ||a^n||^{1/n} = ||a||$$

whenever $a$ is self-adjoint, and thus $\Gamma$ preserves the norms of the self-adjoint elements of $A$. Therefore for any $a \in A$,

$$||\Gamma(a)||^2 = ||\Gamma(a)|| \cdot ||\Gamma(a)|| \quad \text{(C$^*$ identity)}$$
\[ = \| \Gamma(a^*a) \| \quad (\Gamma \text{ is multiplicative}) \]
\[ = \| a^*a \| \quad (\Gamma \text{ is isometric on self-adjoint elements}) \]
\[ = \| a \|^2 , \]
so \( \Gamma \) is isometric.

**Corollary** If \( b \) is a normal element of a C*-algebra, then \( \| b \| = r(b) \).

**Exercises**

1. Give an example of an operator \( T \) on a Hilbert space \( H \) for which the spectral radius \( r(T) \) is strictly less than \( \| T \| \).

2. Prove that if \( b \) is a normal element of a C*-algebra, then \( \| b \| = r(b) \).

3. Let \( \{a_n\} \) be a sequence which converges to 0, and let \( T \) be the weighted shift \( T_x = \sum a_{n+1} \langle x, e_n \rangle e_{n+1} \). What is the spectral radius of \( T \)? The spectral radius can be computed directly, but can be deduced quickly from the spectral properties of compact operators to be developed later.

---

**Part III. The Spectral Theorem. A Continuous Functional Calculus.**

**13. The Spectral Theorem**

From Part II we have tools to study commutative C*-algebras, but the situation we really want to investigate is \( \mathcal{B}(H) \), which is certainly not commutative. What we will do is to start with a normal operator \( T \), and consider the C*-algebra generated by \( T \) (which you can take as the intersection of all C*-subalgebras containing \( I \) and \( T \), or as the closure in \( \mathcal{B}(H) \) of the polynomials in \( T \) and \( T^* \). Take your pick). It’s going to be commutative, and we’ll study it. But perhaps the spectral theory depends on the algebra? Is it possible that \( T \) is invertible in \( \mathcal{B}(H) \) but not in the smaller
C*-algebra generated by T? Before specializing to the study of operators in $\mathcal{B}(H)$ we settle this question and do a little general spectral theory in a Banach algebra.

Now for $a$ in a C*-algebra, denote by $C^*(a)$ the C* subalgebra generated by $a$.

**Proposition** Let $A$ be a C*-algebra with identity $I$, and suppose $a \in A$ has an inverse $a^{-1}$ in $A$. Then $a^{-1} \in C^*(a)$.

**Proof:** Suppose first that $a = a^*$, and let $B$ be the C*-algebra generated by $a$ and $a^{-1}$. Then $B$ is commutative, so the Gelfand transform $\Gamma$ is an isometric *-isomorphism from $B$ onto $C(\chi(B))$. Since $a$ is self-adjoint, $K = \sigma_B(a) \subset \mathbb{R}$, and since $a^{-1} \in B$, $0 \notin K$. Thus the function $h(\lambda) = 1/\lambda$ is continuous on $K$, so by the Weierstrass approximation theorem, there is a sequence of polynomials $\{p_n\}$ which converges to $h$ uniformly on $K$. Notice that if $p(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$, we may define $p(a) = \alpha_0 I + \alpha_1 a + \ldots + \alpha_n a^n$, and for any character $\phi$, $\phi(p(a)) = \alpha_0 + \alpha_1 \phi(a) + \ldots + \alpha_n \phi(a^n)$. Thus for any character $\phi$,

\[
\left(\Gamma(p_n(a))\right)(\phi) - \left(\Gamma(a^{-1})\right)(\phi) = \phi(p_n(a)) - \phi(a^{-1}) = p_n(\phi(a)) - (\phi(a))^{-1} \]

Now $\phi(a) \in K$, and $p_n \rightarrow p$ uniformly on $K$, we obtain (since $\Gamma$ is isometric) $p_n(a) \rightarrow a^{-1}$ in $A$. It follows that when $a$ is self-adjoint, $a^{-1} \in C^*(a)$.

As for the general case, if $a^{-1}$ exists in $A$, then $aa^*$ is self-adjoint and invertible in $A$. Now $a^{-1} = a^*(aa^*)^{-1}$. But $(aa^*)^{-1} \in C^*(aa^*) \subset C^*(a)$. Since $a^* \in C^*(a)$, we have $a^{-1} \in C^*(a)$.

**Corollary** (Spectral Permanence) Let $A$ be a C*-algebra with identity $I$, and let $B$ be a C*-subalgebra of $A$ which contains $I$. Then $\forall b \in B$, $b$ is invertible in $B \iff b$ is invertible in $A$. In particular, for each $b \in B$, $\sigma_B(b) = \sigma_A(b)$.
We are now in a position to give the first of many versions of the spectral theorem.

**Spectral Theorem** Let a be a normal element of a C*-algebra with identity. Let \( A = \mathbb{C}^*(a) \), and let \( \Gamma \) be the Gelfand transform of \( A \). Then

a. \( \Gamma(a) \) is a homeomorphism from \( \chi(A) \) onto \( \sigma(a) \),

b. For each \( f \in C(\sigma(a)) \) there corresponds a unique element \( a_f \in A \) such that \( \Gamma(a_f) = f(\Gamma(a)) \),

c. The map \( \rho: f \rightarrow a_f \) is an isometric *-isomorphism of \( C(\sigma(a)) \) onto \( A \).

The constant function 1 in \( C(\sigma(a)) \) is mapped to \( I \), and the identity function \( \text{id}(\lambda) = \lambda \) is mapped to \( a \in A \).

**proof:** We have already seen that the range of \( \Gamma(a) \) is \( \sigma(a) \). It is not hard to see that if \( \Gamma(a)(\phi_1) = \Gamma(a)(\phi_2) \), then \( \phi_1 \) and \( \phi_2 \) agree on every polynomial in \( a \) and \( a^* \). Since these polynomials are dense in \( A \), \( \Gamma(a) \) is 1-1. Now \( \chi(A) \) is compact and \( \sigma(a) \) is Hausdorff, so \( \Gamma(a) \) is a homeomorphism.

Now \( f \circ \Gamma(a) \) is a composition of continuous functions, and is therefore continuous, so there is a unique element \( a_f \in A \) with \( \Gamma(a_f) = f(\Gamma(a)) \).

Next, \( \Gamma(a) = \text{id} \circ \Gamma(a) \), so by the uniqueness of \( a_{\text{id}} \), \( a = a_{\text{id}} \). We leave it to the reader to check that the map \( f \rightarrow a_f = \Gamma^{-1}(f \circ \Gamma(a)) \) is an isometric *-isomorphism.

Usually \( a_f \) is denoted by \( f(a) \). We shall also use the notation \( \rho(f) \) for \( a_f \).

We'll rephrase this in terms of a normal operator \( T \in \mathcal{B}(H) \) shortly. The main point is that for any function \( f \) which is continuous on \( \sigma(T) \) we can define an operator \( f(T) \). We can clearly do this for a polynomial \( f \), but we've now extended the definition to continuous \( f \). Furthermore, saying that this assignment is an isometric *-isomorphism says that properties enjoyed by the continuous functions are enjoyed by the functions of \( T \). Of course here we mean *-algebra properties. Green-eyed lions need not correspond to green-eyed lions. As an example, if \( f(\lambda) = \overline{f(\lambda)} \) for \( \lambda \in \sigma(T) \), then \( f(T) = f(T)^* \).

We'll make some more comments later, but now let's state the theorem. No proof is necessary, we've already done it.

**Spectral Theorem** Let \( T \) be a normal operator in \( \mathcal{B}(H) \). There is an isometric *-isomorphism \( \rho: C(\sigma(T)) \rightarrow \mathbb{C}^*(T) \), the C*-algebra generated by \( T \). The map \( \rho \) satisfies \( \rho(1) = I \), and \( \rho(\text{id}) = T \).
As we commented above, properties enjoyed by functions on $\sigma(T)$ are enjoyed by the operators $f(T) = \rho(f)$. Implicit in this is the Spectral Mapping theorem.

**Theorem (Spectral Mapping Theorem)** Let $T \in B(H)$ be normal, and let $f \in C(\sigma(T))$. Then $\sigma(f(T)) = f(\sigma(T))$.

**proof:** Since properties of $g \in C(\sigma(T))$ carry over to properties of $g(T)$, $f(T) - \lambda_0 I$ is invertible in $B(H) \iff f(\lambda) - \lambda_0$ is invertible in $C(\sigma(T)) \iff f(\lambda) \cdot \lambda_0 \quad \forall \lambda \in \sigma(T)$.

### 14. Some Applications of the Spectral Theorem

In this section we use the Spectral Theorem to obtain information about a normal operator.

**Theorem** Let $T$ be a normal operator in $B(H)$. Then

- a. $T$ is self-adjoint $\iff \sigma(T) \subset \mathbb{R}$,
- b. $T$ is unitary $\iff \sigma(T) \subset T = \{\lambda : |\lambda| = 1\}$.

**proof:** The proofs of the $\Rightarrow$ implications (if $T$ is self-adjoint, then $\sigma(T) \subset \mathbb{R}$, and if $T$ is unitary, $\sigma(T) \subset T$) do not use the spectral theorem and we leave these proofs to the reader. In the other direction, we use the Spectral Theorem and the map $\rho: C(\sigma(T)) \rightarrow B(H)$. If $\sigma(T) \subset \mathbb{R}$, then the functions $f(\lambda) = \lambda$ and $g(\lambda) = \overline{\lambda}$ agree on $\sigma(T)$. Therefore the operators $f(T) = T$ and $g(T) = T^*$ are the same. If $T$ is normal and $\sigma(T) \subset T$, then the functions $f(\lambda) = \overline{\lambda}$ and $g(\lambda) = \lambda$ satisfy $fg = gf = 1$ on $\sigma(T)$. Therefore $T^*T = TT^* = I$, and $T$ is unitary.

The preceding result remains true in any C*-algebra. Portions of the next result make sense in a general C*-algebra, and are true there. We say an element $a \in A$ **positive** if $a = b^*b$ for some $b \in A$. We write $a \cdot 0$, and notice that $a$ is necessarily self-adjoint.

**Theorem** Let $T$ be a normal operator on a Hilbert space $H$. Then $T \cdot 0 \iff \langle Tx, x \rangle \cdot 0$ for all $x \in H \iff \sigma(T) \subset [0, \cdot) \subset \mathbb{R}$.

We leave the proof to the reader. Please note which implications hold in any C*-algebra, and which are valid without the assumption of normality. Note
also that we can define an order in $\mathcal{B}(H)$ by $T \cdot S \iff (T - S) \cdot 0$, and that $f(\lambda) \cdot 0$ in $C(\sigma(T)) \iff f(T) \cdot 0$ in $\mathcal{B}(H)$.

We’ll give a couple of applications involving positive operators. Again, please see which make sense in any C*-algebra.

**Theorem.** Let $T$ be a positive operator in $\mathcal{B}(H)$. Then there is a unique positive operator $S$, usually denoted by $\sqrt{T}$ such that $S^2 = T$.

**proof:** $T$ is positive, and thus $\sigma(T) \subseteq [0, \cdot)$, so the function $f(\lambda) = \sqrt{\lambda}$ is well defined, positive on $\sigma(T)$, and $f(\lambda)^2 = \lambda$ on $\sigma(T)$. Therefore $\sqrt{T} = f(T)$ is positive and satisfies $\sqrt{T}^2 = T$. We leave the uniqueness as an exercise.

In a similar manner, one may define $|T|$ for any $T \in \mathcal{B}(H)$, and in fact $|T| = T^* T$.

We saw earlier that each element of a *-algebra may be written in a form analogous to $z = \text{Re}(z) + i \text{Im}(z)$ for complex numbers. Each complex number $z$ may also be written in polar form $z = r e^{i \theta}$. The analogous statement for operators would be $T = P U$ with $P \cdot 0$ and $U$ unitary, and it’s not true. The problem is with $U$ (try right shift), and we’ll do as well as one can.

We say an operator $W$ is a **partial isometry** if $\| Wx \| = \| x \|$ for every $x \in \ker W \perp$. That is, $W$ is isometric from its **initial space** $(\ker W) \perp$ to its **final space** $\mathcal{R}(W)$. Notice that $\mathcal{R}(W)$ is necessarily closed.

**Theorem** Every operator $T \in \mathcal{B}(H)$ may be written as $T = W |T|$ where $W$ is a partial isometry with initial space $(\ker T) \perp$ and final space $\overline{\mathcal{R}(T)}$.

**proof:** The representation above is also unique in some sense. We shall let the reader formulate and prove the uniqueness statement.

Our first observation is useful in many contexts: $\ker T = \ker (T^* T)$. One inclusion is clear, and if $x \in \ker (T^* T)$, then $0 = \langle T^* T x, x \rangle = \langle T x, T x \rangle = \| T x \|^2$, so $x \in \ker T$. But we get more out of this, namely $\| |T| x \|^2 = \| T x \|^2$ for every $x \in H$. Now $W$, defined on $\mathcal{R}(|T|)$ has to map a vector $|T| x$ to $T x$, and we have shown that $W$ is isometric. Since it also shows $\ker |T| = \ker T$, we may as well assume $x \in (\ker T) \perp$, from which we see that

School of Mathematics
Georgia Institute of Technology
W is linear. It remains only to show that W maps \((\ker T)^\perp\) to \(\mathcal{R}(T)\). At present we have W defined on \(\mathcal{R}(|T|)\), so we wish to show \(\mathcal{R}(|T|)\) is dense in \((\ker T)^\perp\). But we have shown that \(\ker(T^*T) = \ker(T) = \ker(|T|)\). Thus
\[
\mathcal{R}(|T|) = (\ker|T^*|)^\perp = (\ker|T|)^\perp = (\ker T)^\perp.
\]

As a final application for this section, we examine eigenvalues of a normal operator. We'll be able to do better than this after we extend our functional calculus from continuous functions to bounded measurable functions on \(\sigma(T)\).

**Theorem** Let \(\lambda_0\) be an isolated point in the spectrum of a normal operator \(T\). Then \(\lambda_0 \in \sigma_p(T)\), i.e., \(\lambda_{\text{Error!}}\)

**proof:** Let \(f\) be a function which is continuous on \(\sigma(T)\) and assumes the value 1 at \(\lambda = \lambda_0\) and the value 0 on \(\sigma(T) \setminus \{\lambda_0\}\). Since \(f\) is real valued and \(f^2 = f\), \(f(T)\) is an orthogonal projection. Now \((\lambda_{\text{Error!}}) f(T)\) is the 0 operator since \(\| (\lambda_{\text{Error!}}) f(\lambda) \|_{\text{Error!}} = 0\). Since \(f(T) \cdot 0\), we may choose \(y = f(T)x \cdot 0\), and see that \((\lambda_{\text{Error!}}) y = (\lambda_{\text{Error!}}) f(T) x = 0\), so \(\lambda_{\text{Error!}}\)

Looking more closely at this argument, we see that \(\mathcal{R}(f(T)) \subset \text{EigenSpace}_{\lambda_{\text{Error!}}}\) be as in the last paragraph, and let \(g(\lambda) = 1 - f(\lambda)\). Let
\[
h(\lambda) = \begin{cases} 1/(\lambda - \lambda_0) & \text{if } \lambda \cdot \lambda_0 \\ 0 & \text{otherwise} \end{cases}.
\]
Then we have \((\lambda - \lambda_0) h(\lambda) = g(\lambda)\), so \((T - \lambda_0 I) h(T) = h(T)(T - \lambda_0 I) = g(T)\). Thus if \(Tx = \lambda_0 x\), we have \(g(T)x = 0\). But \(g(\lambda) + f(\lambda) = 1\), so \(g(T)x + f(T)x = x\). It follows that \(f(T)x = x\), so \(\text{EigenSpace}_{\lambda_0} \subset \mathcal{R}(f(T))\). \(f(T)\) then is the orthogonal projection onto \(\text{EigenSpace}_{\lambda_0}\).

**Exercises**

1. Let \(T\) be a bounded operator on a Hilbert space. Show that
   a. If \(T\) is self-adjoint, then \(\sigma(T) \subset \mathbb{R}\). (Don’t repeat the proof in Section 12.)
   b. If \(T\) is unitary, then \(\sigma(T) \subset T = \{ \lambda : |\lambda| = 1 \}\).

2. Recall that an operator \(T \in B(H)\) is positive \((T \cdot 0)\) if \(T = A^*A\) for some \(A \in B(H)\). Show that \(T \cdot 0 \Leftrightarrow \text{sp}(T) \subset [0, \cdot) \Leftrightarrow \langle Tx, x \rangle \cdot 0\) for all \(x \in H\). Certain implications require that \(T\) be normal. Note which these are.

3. Prove the uniqueness of the positive square root. State and prove a uniqueness result for the polar decomposition.
4. Compute the polar decomposition of the following operators.

a. The right shift $S$ on $l_2(\mathbb{N})$.

b. The left shift $L$ on $l_2(\mathbb{N})$.

c. The right shift $S$ on $l_2(\mathbb{Z})$.

d. The multiplication operator $M_\phi$ on $l_2(\mathbb{N})$ defined by $M_\phi e_i = \phi(i)e_i$.

We denote the unit vectors by $e_i$, and assume the function $\phi$ is bounded.

e. The weighted shift $S M_\phi$ on $l_2(\mathbb{N})$.

5. Show that if $W$ is a partial isometry, then $W^*W$ is a projection onto the initial space of $W$, and that $WW^*$ is a projection onto the final space of $W$.

6. Show that $W \in \mathcal{B}(H)$ is a partial isometry $\iff W^*$ is a partial isometry $\iff W^*W$ is a projection $\iff WW^*$ is a projection $\iff WW^*W = W \iff W^*W = W^*$.

15. Another version of the Spectral Theorem

We now prove another version of the spectral theorem. We show, among other things, that every normal operator on any Hilbert space $H$ can be realized as a multiplication operator on an appropriate $L^2(\mu)$. The result is interesting and useful in its own right, and the proof will give an idea of arguments yet to come.

Let $A$ be a C*-subalgebra of $\mathcal{B}(H)$. We say that $x \in H$ is a *-cyclic vector for $A$ if $\{Tx : T \in A\}$ is dense in $H$. We say $x$ is a *-cyclic vector for a normal operator $T$ if it is *-cyclic for $C^*(T)$.

There are a couple of ways to understand what it means to be a *-cyclic vector, and to this end we state a simple proposition. A subspace satisfying either condition of this proposition is sometimes called a reducing subspace for $T$.

**Proposition** Let $M$ be a closed subspace of $H$, and let $T \in \mathcal{B}(H)$. The following statements are equivalent

a. $T(M) \subseteq M$ and $T(M^\perp) \subseteq M^\perp$
b. \( T(M) \subset M \) and \( T^*(M) \subset M \).

Now a vector \( x \) is *-cyclic for \( T \) if \( \{ p(T,T^*) x : p \text{ is a polynomial} \} = H \), and this is equivalent to saying that the smallest closed subspace of \( H \) containing \( x \) and invariant under \( T \) and \( T^* \) is \( H \). That is, \( x \) is a *-cyclic vector for \( T \) if the only non-trivial reducing subspace for \( T \) containing \( x \) is \( H \).

As an example, let \( H = L^2[0,1] \), and let \( T \) be the multiplication operator \( M_t \). I.e. \( f(t) \to t f(t) \). Then the function \( x(t) \equiv 1 \) is a *-cyclic vector for \( T \), because if \( p \) is a polynomial in two variables, then \( p(T,T^*)x = p(t, \overline{t}) \), and the polynomials in \( t \) and \( \overline{t} \) are dense in \( L^2[0,1] \).

**Theorem** Let \( T \) be a normal operator on \( H \) which has a *-cyclic vector \( x_0 \), and let \( \rho: C(\sigma(T)) \to C^*(T) \) be as in the Spectral Theorem in Section 13. There is a probability measure \( \mu \) on \( \sigma(T) \) and an isometry \( U: L^2(\mu) \to H \) such that \( U^{-1} \rho(f)U = M_f \) for every \( f \in L^2(\mu) \).

**proof:** We may assume \( \| x_0 \| = 1 \). Define a linear functional \( \psi \) on \( C(\sigma(T)) \) by \( \psi(f) = \langle \rho(f) x_0, x_0 \rangle \). Then \( |\psi(f)| \cdot \| \rho(f) x_0 \| \cdot \| x_0 \| \cdot \| f \| \) is bounded. In fact, \( |\psi(1)| = \| \rho(x_0) x_0 \rangle = \| x_0 \| \) is bounded. Furthermore if \( f \in C(\sigma(T)) \) and \( f \cdot 0 \), then \( \rho(f) \cdot 0 \), so \( \psi(f) = \langle \rho(f) x_0, x_0 \rangle = 0 \). By the Riesz Representation Theorem, there is a probability measure \( \mu \) on \( \sigma(T) \) such that \( \psi(f) = \langle \rho(f) x_0, x_0 \rangle = \int f \, d\mu \). Now define \( U: C(\sigma(T)) \to H \) by \( Uf = \rho(f) x_0 \). We shall show that \( U \) is an isometry (when \( f \) is given the \( L^2 \) norm).

It maps \( C(\sigma(T)) \) onto a dense subspace of \( H \) since \( x_0 \) is a cyclic vector for \( T \). Therefore \( U \) extends to an isometry from \( L^2(\mu) \) onto \( H \). Furthermore, for any \( f \in C(\sigma(T)) \), and any \( g \in L^2(\mu) \), \( (U^{-1} \rho(f) U) g = U^{-1} \rho(f) \rho(g) x_0 = U^{-1} \rho(fg) x_0 = f g = M_f g \), as desired.

It remains only to show that \( U \) is isometric. Now \( \| Uf \|^2 = \| \rho(f) x_0 \|^2 = \langle \rho(f) x_0, \rho(f) x_0 \rangle = \langle \rho(\overline{f} f) x_0, x_0 \rangle = \int \overline{f} f \, d\mu = \| f \|^2 \).

We now extend this result from normal operators with a *-cyclic vector to a general normal operator. Unfortunately, the \( L^2 \) we arrive at will not be as intuitive as it was in this case. First, we define the direct sum of a family \( \{H_\gamma\} \) of Hilbert spaces. It is simply
\[ \sum_{H} H_{\gamma} = \{ (x_{\gamma}) : x_{\gamma} \in H_{\gamma} \text{ and } \sum \|x_{\gamma}\|^2 < \cdot \}. \]

It is clear how the inner product and norm should be defined, but keep in mind that the index set need not be countable. A quick application of Zorn’s lemma proves the next proposition.

**Proposition** Let \( T \) be a normal operator on Hilbert space \( H \). Then there is a family of reducing subspaces \( H_{\gamma} \), each of which contains a \(*\)-cyclic vector \( x_{\gamma} \in H_{\gamma} \), and \( H = \sum H_{\gamma} \).

**Theorem** If \( T \) is a normal operator on a Hilbert space \( H \), then there exists a topological space \( X \), a measure \( \mu \) and an isometry \( U : L^2(X, \mu) \to H \) such that \( U^{-1} TU \) is a multiplication operator on \( L^2(X, \mu) \).

**proof:** Write \( H = \sum H_{\gamma} \) as in the proposition, let \( T_{\gamma} = T|H_{\gamma} \), and construct measures \( \mu_{\gamma} \) on \( \sigma(T_{\gamma}) \) and isometries \( U_{\gamma} \) as in the last theorem. Let \( f_{\gamma}(\lambda) = \lambda \) on \( \sigma(T_{\gamma}) \), and let \( (X, \mu) \) be the disjoint union of \( (\sigma(T_{\gamma}), \mu_{\gamma}) \). Define \( f \) on \( X \) by way of \( f_{\gamma} \), and define \( U : \sum H L^2(\sigma(T_{\gamma}), \mu_{\gamma}) \to H \) in the obvious way. Now check the details.

**Exercises**

1. Prove the proposition on reducing subspaces.
2. Complete the details of the last theorem of this section.
3. Let \( T \) be the right shift on \( l^2(\mathbb{Z}) \). Find a cyclic vector for \( T \) and represent \( T \) as a multiplication operator on \( L^2(\mu) \).

**Part IV. The Spectral Theorem. A Functional Calculus for Bounded Measurable Functions.**
16. Resolutions of the Identity

A **representation** of a C*-algebra $A$ is a *-isomorphism from $A$ into the bounded operators on some Hilbert space. In the last part, given a normal operator $T \in \mathcal{B}(H)$, we obtained a representation $\rho: \mathcal{C}(\sigma(T)) \to \mathcal{B}(H)$ which carried the identity function $\text{id}(\lambda) = \lambda$ to $T$. We can thus speak of $f(T)$ whenever $f$ is continuous on $\sigma(T)$. We now plan to extend this to a representation of the C*-algebra of bounded Borel measurable functions on $\sigma(T)$. In the process we'll obtain an expression for $T$ much the same as the expression $A = \sum_{\lambda \in \sigma(A)} \lambda P_k$ for a normal matrix. The sum will be replaced by an integral over $\sigma(T)$, with respect to a projection valued measure (which we shall refer to as a resolution of the identity).

In the following definition, think of the example with $X = [0,1]$, $\Omega = \text{Borel subsets of } [0,1]$, $H = L^2[0,1]$, and $E(\omega)$ the operator on $L^2[0,1]$ given by multiplication by $\chi_{\text{Error!}}$

Let $X$ be a set, $\Omega$ a \(\sigma\)-algebra of subsets of $X$, and let $H$ be a Hilbert space. A **resolution of the identity** for $(X,\Omega,H)$ is a function $E: \Omega \to \mathcal{B}(H)$ satisfying

a. $\forall \omega \in \Omega$, $E(\omega)$ is a self-adjoint projection.

b. $E(\emptyset) = 0$, and $E(X) = I$.

c. $E(\omega_1 \cap \omega_2) = E(\omega_1) E(\omega_2) \forall \omega \in \Omega$.

d. If $\{\omega_k\}$ is a sequence of disjoint sets from $\Omega$, then

$$E\left(\bigcup \omega_k\right) = \sum_{k=1}^{\infty} E(\omega_k).$$

Some comments are in order. First, from c we see that the $E(\omega)$ are commuting projections. Next, if $\omega_1 \cap \omega_2 = \emptyset$, then $E(\omega_1) E(\omega_2) = 0$, so $E(\omega_1)$ and $E(\omega_2)$ have orthogonal ranges. Finally, the series in d. usually has no hope of converging in the norm topology of $\mathcal{B}(H)$. Since the norm of a self-adjoint projection is either 0 or 1, the series $\sum_{k=1}^{\infty} E(\omega_k)$ has little chance of being norm Cauchy. However, by Bessel’s inequality, for any $x \in H$,
\[ \sum_{k=1}^{\cdot} \left\| E(\omega_k)x \right\|^2 \cdot \left\| x \right\|^2 \]. Thus for each \( x \in H \), the series \[ \sum_{k=1}^{\cdot} E(\omega_k)x \]
converges to \( E(\bigcup \omega_k)x \) in the norm topology of \( H \).

Two other useful topologies on \( B(H) \) are the **strong operator topology** and the **weak operator topology**. We say a net \( \{T_\gamma\} \) converges to \( T \) strong operator if for every \( x \in H \), \( T_\gamma x \to Tx \) in the norm topology on \( H \) (i.e., \( \| T_\gamma x - Tx \| \to 0 \) for every \( x \)). We say \( \{T_\gamma x\} \) converges to \( T \) weak operator if for every \( x \in H \), \( \langle T_\gamma x, y \rangle \to \langle Tx, y \rangle \). That is, for any \( x, y \in H \), \( \langle T_\gamma x, y \rangle \to \langle Tx, y \rangle \). What we have just shown is that the series \[ \sum_{k=1}^{\cdot} E(\omega_k) \] converges to \( E(\bigcup \omega_k) \) in the strong operator topology.

A resolution of the identity gives rise to a large family of measures.  

**Proposition** If \( E \) is a resolution of the identity on \( (X, \Omega, H) \), then for each \( x, y \in H \),

\[ E_{x,y}(\omega) = \langle E(\omega)x, y \rangle \]
defines a complex, countably additive measure on \( (X, \Omega) \). It satisfies

\[ \| E_{x,y} \| \cdot \| x \| \cdot \| y \|. \]

Furthermore \( E_{x,x} \) is a positive measure, and \( \| E_{x,x} \| = E_{x,x}(X) = \| x \|^2 \).

**proof:** Showing that \( E_{x,y} \) is a measure is straightforward. We now estimate the total variation \( \| E_{x,y} \| \). Let \( \omega_1, \ldots, \omega_k \) be disjoint measurable sets, and choose \( \theta_j \) so that

\[ \sum_{j=1}^{k} |E_{x,y}(\omega_j)| = \sum_{j=1}^{k} e^{i\theta_j}E_{x,y}(\omega_j) \]

School of Mathematics  
Georgia Institute of Technology
\[
\sum_{j=1}^{k} e^{i\theta_j} \cdot E(\omega_j) \cdot x, y
\]

\[
\sum_{j=1}^{k} E(\omega_j) e^{i\theta_j} \cdot x, y
\]

\[
\| \sum_{j=1}^{k} E(\omega_j) e^{i\theta_j} \cdot x \| \| y \|
\]

\[
\| x \| \| y \| \quad (\text{since } E(\omega_j) \text{ have orthogonal ranges})
\]

This shows \( \| E_{x,y} \| \cdot \| x \| \| y \| \). Since \( E(\omega) \) a self-adjoint projection,

\[
E_{x,x}(\omega) = \langle E(\omega)x, x \rangle = \langle E(\omega)x, E(\omega)x \rangle = \| E(\omega)x \|^2.
\]

Thus we see that \( E_{x,x} \) is a positive measure, and setting \( \omega = X \), we obtain \( \| E_{x,x} \| = \| x \|^2 \).

Now that we have a measure, the urge to integrate is overwhelming. The next proposition explains what we mean by integration of a bounded measurable function with respect to a resolution of the identity. The integral of a simple function \( f = \sum f(t_j) \chi_{\omega_j} \) just has to be \( \sum f(t_j) E(\omega_j) \), so part b. of the following proposition just says that we may integrate \( f \) by approximating it by simple functions. We'll actually come up with the integral by way of part a., but we'll find both this abstract formulation and approximation by simple functions to be useful.

**Theorem** Let \( E \) be a resolution of the identity on \( (X, \Omega, H) \), and let \( f \) be a bounded measurable function on \( X \). Then there is a unique operator \( T \in \mathcal{B}(H) \), which deserves to be called \( \int f(\lambda) dE(\lambda) \) because

a. \( \langle \left( \int f(\lambda) dE(\lambda) \right) x, y \rangle = \int f dE_{x,y} \) for every \( x, y \in H \)

b. If \( \{\omega_1, ..., \omega_k\} \) is a partition of \( X \) such that \( t, t' \in \omega_j \Rightarrow |f(t) - f(t')| < \varepsilon \), then \( \| T - \sum f(t_j) E(\omega_j) \| < \varepsilon \), whenever \( t_j \in \omega_j \).

proof: For \( x, y \in H \), define \( \psi(x, y) = \int f dE_{x,y} \). Then \( \psi \) is a sesquilinear functional on \( H \), and since
\[ |\psi(x,y)| \cdot ||f|| \cdot \mathbb{E}_{x,y} \leq ||f|| \cdot ||x|| \cdot ||y||, \]

\(\psi\) is bounded. Thus there is a unique operator \(T\) which satisfies \(\psi(x,y) = \langle Tx, y \rangle\) for all \(x,y \in H\). This establishes the existence, uniqueness, and the formula in part a. As for part b.,

\[ |\langle T - \sum f(t_j) \mathbb{E}(\omega_j) \rangle x, y \rangle| = |\langle Tx, y \rangle - \sum f(t_j) \langle \mathbb{E}(\omega_j) x, y \rangle| \]

\[ = \left| \sum_{\omega_j} \int f \, d\mathbb{E}_{x,y} - \sum_{j} f(t_j) \langle \mathbb{E}(\omega_j) x, y \rangle \right| \]

\[ = \left| \int_{\omega_j} \int |f(t) - f(t_j)| \, d\mathbb{E}_{x,y} \right| \cdot \epsilon \cdot ||x|| \cdot ||y||. \]

Since \(x\) and \(y\) were arbitrary, part b. is established.

**Exercises**

1. Let \(\{T_n\}\) and \(\{S_n\}\) be sequences in \(B(H)\) which converge to \(T\) and \(S\), respectively, in the strong operator topology. Prove that \(\{S_n \cdot T_n\}\) converges to \(ST\) in the strong operator topology.

2. Let \(\{T_n\}\) and \(\{S_n\}\) be sequences in \(B(H)\) and suppose \(S_n \to S\) in the weak operator topology and \(T_n \to T\) in the strong operator topology. Prove that \(\{S_n \cdot T_n\}\) converges to \(ST\) in the weak operator topology.

3. Prove that composition in \(B(H)\) is separately continuous but not jointly continuous in the weak operator topology.

4. Let \(\{e_n\}\) be a complete orthonormal sequence for a separable Hilbert space \(H\). Show that the weak closure of \(W = \{\sqrt{n} \cdot e_n\}\) contains 0. Thus there is a net \(\{\sqrt{n}_\gamma \cdot e_{n_\gamma}\}\) from \(W\) which converges weakly to zero. Is there a sequence from \(W\) which converges weakly to 0? What sort of
directed set does the net \( \{\sqrt{n_{\gamma}} e_{n_{\gamma}} \} \) have? Define \( T_{\gamma} x = \sqrt{n_{\gamma}} \cdot x, e_{n_{\gamma}} \cdot e_{n_{\gamma}} \), and show that \( T_{\gamma} \rightarrow 0 \) strong operator but \( T_{\gamma}^2 \) does not converge to 0 strong operator.

5. Explain why the functional \( \psi \) in the proof of the last theorem is sesquilinear.

17. Representations and Resolutions of the Identity

Here’s the plot for this section. Using the last Theorem, we’ll be able to show how every resolution of the identity gives rise to a representation \( \rho \) of the space \( B(X) \) of bounded measurable functions on \( X \). Next we’ll see that each representation of \( C(X) \) corresponds to a resolution of the identity. But we know that if we start with a normal operator \( T \), we have a representation \( \rho: C(\sigma(T)) \rightarrow B(H) \). Thus to each normal operator \( T \) there will correspond a (unique) representation \( \rho \) of the bounded measurable functions on \( \sigma(T) \). This representation extends that from \( C(\sigma(T)) \rightarrow B(H) \), and we will obtain a spectral theorem with a measurable functional calculus.

A word should be said about the space \( B(X) \) and its norm. As usual, we won’t make the distinction between functions which differ on a set of measure zero, but what in this context constitutes a set of measure zero? The norm is the sup norm, or more precisely stated, the essential sup norm. To define this, let \( E \) be a resolution of the identity on \( (X, \Omega, H) \), and let \( \{D_i\} \) be a countable collection of disks which is a basis for the topology for \( C \). For \( f \in B(X) \), let \( \mathcal{U}_f = \bigcup \{ D_i : E(f^{-1}(D_i)) = 0 \} \). This is an open set, and the closed set \( C \setminus \mathcal{U}_f \) is the essential range of \( f \). The norm \( \| f \| = \| f \|_{\bullet} \) is understood to be the essential sup of \( f \), which is \( \sup \{ |\lambda| : \lambda \in \text{essential range of } f \} \).

**Theorem** If \( E \) is a resolution of the identity on \( (X, \Omega, H) \), and \( \rho: B(X) \rightarrow B(H) \) is defined by \( \rho(f) = \int \! f dE \), then \( \rho \) is an isometric representation of \( B(X) \). For any \( x, y \in H \), we have \( \langle \rho(f)x, y \rangle = \int \! f dE_{x,y} \), and \( \| \rho(f)x \|^2 = \int |f|^2 dE_{x,x} \).

Furthermore, \( S \in B(H) \) commutes with every operator \( \rho(f) \) if and only if \( S \) commutes with every projection \( E(\omega) \).
proof: It is not hard to show that \( \rho \) is linear and that \( \rho(\overline{f}^*) = \rho(f)^* \). Using part b. of the last theorem, we shall prove that for \( f, g \in B(X) \), \( \rho(fg) = \rho(f)\rho(g) \). First, let \( \varepsilon > 0 \), and choose a partition \( \{\omega_1, \ldots, \omega_k\} \) such that \( t, t' \in \omega_j \Rightarrow |h(t) - h(t')| < \varepsilon \) for \( h = f, g, \) or \( fg \). Then

\[
\|\int fg \, dE - \int f \, dE \int g \, dE\| \cdot \varepsilon + \|\sum f(t_j)g(t_j)E(\omega_j) - \left(\sum f(t_j)E(\omega_j)\right) \left(\sum g(t_j)E(\omega_j)\right)\| \cdot \varepsilon
\]

Now the first term is less than \( \varepsilon \), and since the \( E(\omega_j) \) are self-adjoint projections with orthogonal ranges, the middle term is 0. We then obtain

\[
\|\int fg \, dE - \int f \, dE \int g \, dE\| \cdot \varepsilon + \|\sum f(t_j)E(\omega_j) - \int f \, dE \int g \, dE\| \cdot \varepsilon + \|f\| \cdot \varepsilon + \|g\| \cdot \varepsilon ,
\]

and this establishes the multiplicativity of \( \rho \).

That \( \langle \rho(f)x, y \rangle = \int f dE_{x,y} \) follows directly from the definition of \( \rho \), and \( \|\rho(f)x\|^2 = \langle \rho(f)x, \rho(f)x \rangle = \langle \rho(\overline{f}^*)fx, x \rangle = \int |f|^2 dE_{x,x} \). To see that \( \rho \) is isometric, note first that if \( f = \chi_{\omega} \) is a characteristic function and if \( x = E(\omega)x \), then \( \|\rho(f)x\| = \|E(\omega)x\| = \|x\| \), so \( \|\rho(f)\| = \|f\| \). It follows that \( \rho \) is isometric on simple functions, and hence on \( B(X) \) by approximation with simple functions.
As for the commutativity statement, if $S$ commutes with every $\rho(f)$, then certainly $S$ commutes with $\rho(\chi_\omega) = E(\omega)$, and if $S$ commutes with every $E(\omega)$, then approximation by simple functions shows that $S$ commutes with every $\rho(f)$.

**Theorem** Let $\rho: C(X) \to \mathcal{B}(H)$ be a representation. Then there is a unique resolution of the identity $E$ on the Borel sets of $X$ such that $\rho(f) = \int f \, dE$ for every $f \in C(X)$. An operator $S \in \mathcal{B}(H)$ commutes with every $\rho(f)$ if and only if $S$ commutes with every $E(\omega)$.

**proof:** The idea is clear. We'd like $E(\omega) = \int \chi_\omega \, dE$, so we'd like $E(\omega) = \rho(\chi_\omega)$, but $\chi_\omega$ probably isn't continuous so $\rho(\chi_\omega)$ isn't yet defined. The situation is much like that in identifying $C(X)^*$, and we will use the Riesz Representation Theorem for $C(X)$.

For any $x, y \in H$, the map $\tau: f \to \langle \rho(f) \, x, y \rangle$ is a linear functional on $C(X)$, and $\| \tau(f) \| \cdot \| f \| \cdot \| x \| \cdot \| y \|$ so $\tau$ is continuous. Hence there exists a unique measure $\mu_{x,y}$ on the Borel sets of $X$ with the property

$$\langle \rho(f) \, x, y \rangle = \int \mu_{x,y} \, f \, d\mu_{x,y} \quad \forall f \in C(X).$$

Now $\| \mu_{x,y} \| \cdot \| x \| \cdot \| y \|$, and by the uniqueness of $\mu_{x,y}$, the map $(x,y) \to \mu_{x,y}$ is linear in $x$ and conjugate linear in $y$. Now for $f \in B(X)$, define $\psi(x,y) = \int f \, d\mu_{x,y}$. Then $\| \psi(x,y) \| \cdot \| f \| \cdot \| x \| \cdot \| y \|$, so $\psi$ is a bounded sesquilinear function on $H$, and there exists a unique operator $\hat{\rho}(f)$ with

$$\psi(x,y) = \int f \, d\mu_{x,y} = \langle \hat{\rho}(f) \, x, y \rangle.$$ 

We now claim that $\hat{\rho}$ is a representation of $B(X)$ which extends $\rho$. It is clear that $\hat{\rho}$ extends $\rho$, and that $\hat{\rho}$ is linear. We'll show $\hat{\rho}$ is multiplicative in a moment. First let's use it to construct the resolution of the identity. We define $E(\omega) = \rho$

**Error!**

$$\| E(\bigcup \omega_k) \, x - \sum_{k=1}^{N} E(\omega_k) \, x \|^2$$
= \left\langle E(\Omega_N)\ x, \ E(\Omega_N)\ x \right\rangle

= \left\langle E(\Omega_N)\ x, \ x \right\rangle

= \left\langle \hat{\rho}(\chi_{\Omega_N})\ x, \ x \right\rangle

= \int \chi_{\Omega_N} \ d\mu_{x,x}

= \sum_{k=N+1}^{\infty} \mu_{x,x}(\omega_k), \text{ which tends to 0 as } N \to \infty.

By approximation with simple functions it is clear that \( \hat{\rho}(f) = \int f \ dE \quad \forall \ f \in C(X). \)

It remains, then, to show the multiplicativity of \( \hat{\rho} \), and since we know \( \hat{\rho} \) extends \( \rho \), \( \hat{\rho}(fg) = \hat{\rho}(f) \hat{\rho}(g) \) whenever \( f \) and \( g \) are continuous. Now Goldstine's Theorem asserts that for any Banach space \( Z \), the unit ball of \( Z \) is weak* dense in the unit ball of \( Z^{**} \). Therefore, whenever \( f \in B(X) \), we may find a net \( \{f_{\gamma}\} \) in \( C(X) \) which converges to \( f \) weak*. That is, \( \forall \ \mu, \)

\[ \int f_{\gamma} \ d\mu \to \int f \ d\mu. \]

Now using the measure \( gd\mu \), we obtain

\[ \int f_{\gamma}g \ d\mu \to \int fg \ d\mu \quad \forall \ \mu. \]

Thus

\[ \left\langle \hat{\rho} (f,g)\ x, \ y \right\rangle = \int f_{\gamma}g \ d\mu_{x,y} \]

\[ \to \int fg \ d\mu_{x,y}. \]
That is, \( \hat{\rho}(f \gamma g) \to \hat{\rho}(fg) \) in the weak operator topology. Taking \( g \in \mathbb{C}(X) \), we see that whenever \( f \) is bounded measurable and \( g \) is continuous, we have multiplicativity, since (all limits are weak operator)

\[
\hat{\rho}(fg) = \lim \hat{\rho}(f \gamma g) \\
= \lim \rho(f \gamma g) \text{ (since } \hat{\rho} \text{ extends } \rho) \\
= \lim \rho(f \gamma) \rho(g) \\
= \hat{\rho}(f) \rho(g).
\]

Similarly, if \( f \) is continuous and \( g \) is bounded and measurable, \( \hat{\rho}(fg) = \rho(f)\hat{\rho}(g) \). We now have assembled the following facts: \( \{f \gamma\} \) is a net of continuous functions converging weak* to \( f \), so \( \hat{\rho}(f \gamma) \to \hat{\rho}(f) \) and \( \hat{\rho}(f \gamma g) \to \hat{\rho}(fg) \). Thus \( \hat{\rho}(fg) = \lim \hat{\rho}(f \gamma g) = \lim \rho(f \gamma) \hat{\rho}(g) = \hat{\rho}(f) \rho(g) \).

As for the commutativity statement, notice that for every \( x, y \in H \), we have

\[
\langle S \rho(f)x, y \rangle = \int f \, dE_{x,S^*y}, \\
\langle \rho(f)Sx, y \rangle = \int f \, dE_{Sx,y},
\]

and

\[
\langle S E(\omega)x, y \rangle = \langle E(\omega)x, S^*y \rangle = E_{x,S^*y}(\omega), \\
\langle E(\omega)Sx, y \rangle = E_{Sx,y}(\omega).
\]

Thus the measures \( E_{x,S^*y} \) and \( ES_{x,y} \) are equal if and only if \( S E(\omega) = E(\omega)S \) for all \( \omega \) if and only if \( S \rho(f) = \rho(f)S \) for all \( f \).

We now present a version of the Spectral Theorem which represents bounded measurable functions of the operator \( T \).
**Spectral Theorem** Let \( T \) be a normal operator in \( \mathcal{B}(H) \). There exists a unique resolution of the identity \( E \) on the Borel sets of \( \sigma(T) \) such that

a. \( \rho : \mathcal{B}(\sigma(T)) \to \mathcal{B}(H) \) defined by \( \rho(f) = \int \sigma(T) f(\lambda) \, dE(\lambda) \) is an isometric representation which carries 1 to I and \( \text{id}(\lambda) = \lambda \) to \( T \). In particular, \( T = \int \sigma(T) \lambda \, dE(\lambda) \).

b. If \( G \) is a relatively open subset of \( \sigma(T) \), then \( E(G) > 0 \).

c. An operator \( S \) on \( H \) commutes with \( T \) if and only if \( S \) commutes with all the projections \( E(\omega) \).

*proof:* We have already an isometric *-isomorphism \( \rho \) from \( C(\sigma(T)) \) onto \( C^*(T) \). This yields a unique resolution of the identity \( E \) on the Borel sets of \( \sigma(T) \) with \( \rho(f) = \int f \, dE \) whenever \( f \) is continuous on \( \sigma(T) \). In particular, \( \rho(\text{id}) = T = \int \sigma(T) \lambda \, dE(\lambda) \). By the previous results this extends to an isometric representation of \( \mathcal{B}(\sigma(T)) \).

If \( G \) is relatively open, we can find a continuous function \( g \) with \( 0 \leq g \leq \chi_G \). Since a representation preserves positivity, \( E(G) = \rho(\chi_G) \cdot \rho(g) > 0 \).

Since \( \rho \) is an isometry from \( C(\sigma(T)) \) onto \( C^*(T) \), the commutativity statement will follow provided whenever \( S \) commutes with \( T \), then \( S \) commutes with every operator in \( C^*(T) \). Taking \( M = N = T \) in the next theorem, we see that if \( S \) commutes with \( T \), then \( S \) commutes with \( T^* \), and hence with every operator in \( C^*(T) \).

**Theorem** (Fuglede-Putnam-Rosenblum) Suppose \( S \in \mathcal{B}(H) \), that \( M \) and \( N \) are normal operators in \( \mathcal{B}(H) \), and that \( MS = SN \). Then \( M^*S = SN^* \).

*proof:* Notice first that for any operator \( T \),

\[
W = \exp(T - T^*) = \sum_{n=0}^{\infty} \frac{1}{n!} (T - T^*)^n
\]

satisfies \( W^* = W^{-1} \) and is therefore unitary. In particular,
The hypothesis that $MS = SN$ implies that $\exp(M) S = S \exp(N)$, so
$S = \exp(-M) S \exp(N)$. Since $M$ and $N$ are normal, we may deduce that

$$\exp(M^*) S \exp(-N^*) = \exp(M^* - M) S \exp(N - N^*),$$

from which it follows that

$$\| \exp(M^*) S \exp(-N^*) \| \cdot \| S \|.$$

This same argument applies to $f(\lambda) = \exp(\lambda M^*) S \exp(-\lambda N^*)$ and shows that $f(\lambda)$ is a bounded entire function. By Liouville’s Theorem, $f(\lambda)$ is constant, and in particular, $f(\lambda) = f(0) = S$ for all $\lambda$. Thus

$$\exp(\lambda M^*) S = S \exp(\lambda N^*)$$

for all $\lambda$. Expanding these exponentials in series and equating the coefficients of $\lambda$ yields $M^* S = S N^*$.

---

### 18. Some Applications of the Spectral Theorem

We turn to a few applications of the Spectral Theorem proved in Section 17. We begin with a return to eigenvalues.

**Theorem.** Let $T$ be a normal operator. $\lambda_0 \in \sigma(T)$ is an eigenvalue of $T$ if and only if $E(\{\lambda_0\}) \cdot 0$. If $E(\{\lambda_0\}) \cdot 0$, it is the orthogonal projection onto the eigenspace for $\lambda_0$.

**proof:** The proof will look much as in Section 14, where we considered an isolated point in $\sigma(T)$. Assume $E(\{\lambda_0\}) \cdot 0$, and again, let $f(\lambda) = \chi_{\{\lambda_0\}}$. Then

$$(\lambda - \lambda_0) f(\lambda) = 0,$$

but $f(\lambda) \cdot 0$. Choosing $y \in \mathbb{R}(f(T)), y \cdot 0$, we have $(T - \lambda_0) y = 0$, so $\lambda_0$ is an eigenvalue of $T$. In addition, $f(T) = E(\{\lambda_0\})$, so we have shown that $\mathbb{R}(E(\{\lambda_0\})) \subset \ker(T - \lambda_0 I)$. For the other inclusion, let $\omega_n = \{ \lambda \in \\sigma(T) : |\lambda - \lambda_0| > 1/n \}$, and define $g_n(\lambda) = \begin{cases} \frac{1}{\lambda - \lambda_0} & \lambda \in \omega_n \\ 0 & \text{otherwise} \end{cases}$. Then each
$g_n$ is a bounded Borel function on $\sigma(T)$, and $g_n(\lambda) (\lambda - \lambda_0) = \chi_{\omega_n}$ for each $n$. It follows that $g_n(T)(T - \lambda_0 I) = E(\omega_n)$ for each $n$, and if $(T - \lambda_0 I)y = 0$, then $E(\omega_n)y = 0$ for all $n$. By the countable additivity of $E$, $E(\sigma(T) \setminus \{\lambda_0\})y = 0$, so $E(\{\lambda_0\})y = y$. This shows that $\ker(T - \lambda_0 I) \subset \mathbb{R}(E(\{\lambda_0\}))$ and proves that if $\lambda_0$ is an eigenvalue of $T$ then $E(\{\lambda_0\}) \cdot 0$.

The next result shows that the norm of a normal operator is determined by its **numerical range**, $\{\langle Tx, x \rangle : x \in H\}$. It can also be proved without using the spectral theorem. Try it.

**Theorem** Let $T \in \mathcal{B}(H)$ be normal. Then $\|T\| = \sup \{\|\langle Tx, x \rangle\| : x \in H\}$.

**proof:** $\|T\| = \|\text{id}\| = \sup\{|\lambda| : \lambda \in \sigma(T)\}$, so there exists $\lambda_0 \in \sigma(T)$ such that $|\lambda_0| = \|T\|$. If $\lambda_0$ were an eigenvalue, we’d be done, since, choosing an eigenvector $x_0$ of norm one, we’d have $\langle Tx_0, x_0 \rangle = \lambda_0$. Now let $\varepsilon > 0$, and let $\omega = \{\lambda \in \sigma(T) : |\lambda - \lambda_0| < \varepsilon\}$. $\omega$ is relatively open, so $E(\omega) \cdot 0$. Define $f(\lambda) = \begin{cases} (\lambda - \lambda_0) & \lambda \in \omega \\ 0 & \text{otherwise} \end{cases}$. Then $f(\lambda) = (\lambda - \lambda_0)\chi_{\omega}$, so $f(T) = (T - \lambda_0 I)E(\omega)$. Now let $x$ be a unit vector from $\mathbb{R}(E(\omega))$. Then $f(T)x = Tx - \lambda_0 x$, so $|\langle Tx, x \rangle - \lambda_0| = |\langle Tx, x \rangle - \lambda_0 - x, x \rangle| = |\langle Tx - \lambda_0 x, x \rangle| = |\langle f(T)x, x \rangle|$. \hfill $\star \|f(T)\| < \varepsilon$.

**Exercises**

1. Show that if $T$ is a normal operator on a separable Hilbert space $H$, then there are at most countably many points $\lambda$ with $E(\{\lambda\}) \cdot 0$.

2. Let $T$ be a normal operator on a separable Hilbert space $H$. Show that $E(\sigma(T) \setminus \sigma_p(T)) = 0$ if and only if there is a basis for $H$ consisting of eigenvectors for $T$. What if $H$ is not separable?

2. Let $T$ be a self-adjoint operator in $\mathcal{B}(H)$. Prove that there are positive operators $T^+$ and $T^-$ such that $T = T^+ - T^-$. What is the operator $T^+ + T^-$?

3. Without using the spectral theorem, prove that for a normal operator $T$, $\|T\| = \sup \{|\langle Tx, x \rangle| : x \in H\}$. 

School of Mathematics
Georgia Institute of Technology

In this section we present two types of results. First we analyze a compact normal operator in terms of its spectrum, and second we prove that if $H$ is separable, the only proper, closed two-sided, *-stable ideal (in the rest of this section we'll just say ideal) in $\mathcal{B}(H)$ is the ideal of compact operators. Compact operators can be expressed quite well in terms of their spectrum, even when they are on a Banach space rather than Hilbert space. The reader should investigate these matters.

**Theorem** Let $T$ be a normal operator on a Hilbert space $H$, and let $E$ be the resolution of the identity for $T$. $T$ is compact if and only if for every $\varepsilon > 0$, $E(\{\lambda : |\lambda| > \varepsilon\})$ is of finite rank.

**proof:** Suppose that $\forall \varepsilon > 0$, $E(\{\lambda : |\lambda| > \varepsilon\})$ is finite rank. Let $\omega_n = \{\lambda : |\lambda| > 1/n\}$, and let $f_n(\lambda) = \lambda \chi_{\omega_n}(\lambda)$. Then $f_n(T) = \int f_n(\lambda)dE(\lambda)$ is finite rank, and $\|T - f_n(T)\| \cdot \|\text{id}(\lambda) - f_n(\lambda)\| \cdot 1/n$. Thus $T$ is compact, being the norm limit of finite rank operators.

On the other hand, suppose there exists $\varepsilon > 0$ with $\mathcal{R}(E(\{\lambda : |\lambda| > \varepsilon\}))$ infinite dimensional. Let $f(\lambda) = \begin{cases} 1/\lambda, & \lambda \in \omega = \{\lambda : |\lambda| > \varepsilon\} \\ 0, & \text{otherwise} \end{cases}$. Then $\lambda f(\lambda) = \chi_{\omega}$, so $Tf(T) = E(\omega)$. Since $E(\omega)$ has infinite dimensional range it cannot be compact, so neither can $T$.

This result can be used to prove the next theorem about a compact normal operator. We leave the details to the reader.

**Theorem** Let $T$ be a compact normal operator on $H$. Then

a. If $\lambda_0$ is a cluster point of $\sigma(T)$, then $\lambda_0 = 0$. Thus $\sigma(T)$ is countable and each non-zero element of $\sigma(T)$ is an eigenvalue of $T$.

b. For each $\lambda_0 \neq 0$ in $\sigma(T)$, the eigenspace of $\lambda_0$ is finite dimensional.

c. There is an orthonormal basis $\{x_n\}$ for $H$ consisting of eigenvectors for $T$. That is $Tx = \sum l(\lambda_n, x, x_n) \cdot x_n \forall x \in H$.

Now we consider the ideals in $\mathcal{B}(H)$ when $H$ is separable. We'll denote the ideal of compact operators by $\mathcal{C}(H)$, and we'll leave it to the reader to
check that any ideal in \( B(H) \) contains all rank one operators. It thus contains all finite rank operators, and hence all compact operators (assuming it is closed). Therefore if \( J \) is any ideal in \( B(H) \), \( \lambda(H) \subset J \). Now suppose \( J \) contains a normal non-compact operator \( T \), with spectral resolution \( E \). Then there exists \( \varepsilon > 0 \) such that with \( \omega = \{ \lambda : |\lambda| > \varepsilon \} \), we have \( E(\omega) \) is of infinite rank.

Again letting \( f(\lambda) = \frac{1}{\lambda} \chi_\omega \), we see that \( f(T)T = E(\omega) \) belongs to \( J \) and has infinite dimensional range. Now since \( H \) is separable, \( \dim(\mathcal{R}(E(\omega))) = \dim H \), so there is a partial isometry \( W \) with initial space \( (\ker W)^\perp = \mathcal{R}(E(\omega)) \) and final space \( H \). But then \( WE(\omega)W^* = I \) belongs to \( J \). Now in the general case, if \( J \) contains a noncompact operator \( T \), then \( T^*T \) belongs to \( J \), is normal and not compact.

**Exercises**

1. Let \( T \) be a compact normal operator on \( H \). Prove that
   a. If \( \lambda_0 \) is a cluster point of \( \sigma(T) \), then \( \lambda_0 = 0 \). Thus \( \sigma(T) \) is countable and each non-zero element of \( \sigma(T) \) is an eigenvalue of \( T \).
   b. For each \( \lambda_0 \cdot 0 \) in \( \sigma(T) \), the eigenspace of \( \lambda_0 \) is finite dimensional.
   c. There is an orthonormal basis \( \{ x_n \} \) for \( H \) consisting of eigenvectors for \( T \). That is \( Tx = \sum \lambda_n \langle x, x_n \rangle x_n \) \( \forall x \in H \).

2. Prove that if \( T^* \) is compact then \( T \) is compact and conversely.

**Part V. Unbounded Operators.**

**20. Unbounded Operators, Terminology and Examples**

From now on, by an **operator** from a Hilbert space \( H \) to a Hilbert space \( K \), we mean a function \( T \) defined on a subspace \( D(T) \subset H \) and taking values in \( K \), which satisfies \( T(\alpha x + \beta y) = \alpha Tx + \beta Ty \) whenever \( \alpha, \beta \in \mathbb{C} \), and \( x, y \in D(T) \). In particular, when we say "\( T \) is an operator on \( H \)" we do not mean to imply that \( Tx \) is defined for all \( x \in H \) nor that \( T \) is bounded. Notice that if \( T \) is bounded, i. e., \( \| Tx \| \leq M \| x \| \) for some \( M \), then we may extend the definition of \( T \) to \( D(T) \). Generally, the domain of \( T \) will not be closed, and
in most cases, for reasons which will become clear in a moment, we will
require \( D(T) \) to be dense in \( H \). It will become important to keep track of the
domains of operators. For example \( D(S + T) = D(S) \cap D(T) \), and if \( S \) and \( T \)
are both operators on \( H \), then \( D(ST) = \{ x : x \in D(T) \text{ and } Tx \in D(S) \} \).

We say \( T \) is an **extension** of \( S \) if \( D(S) \subseteq D(T) \), and \( Tx = Sx \) for any
\( x \in D(S) \). We shall write \( S \subseteq T \) to indicate that \( T \) is an extension of \( S \). The **graph** of an operator \( T \) is \( G(T) = \{(x,Tx) : x \in D(T) \} \subseteq H \oplus K \), and certainly
\( S \subseteq T \Leftrightarrow G(S) \subseteq G(T) \). Subspaces of \( H \oplus K \) can be recognized as graphs by the
following proposition.

**Proposition** A subspace \( G \) of \( H \oplus K \) is the graph of an operator from \( H \) to \( K \)
if and only if \( y \in K \), and \((0,y) \in G \) implies \( y = 0 \).

We say that an operator \( T \) is **closed** if its graph \( G(T) \) is a closed
subspace of \( H \oplus K \). Notice that if \( D(T) \) is closed and \( T \) is closed, then by the
closed graph theorem, \( T \) is bounded. We say \( T \) is **closeable** if it has a closed
extension. If \( T \) is closeable, then \( G(T) \) satisfies the hypotheses of the last
proposition, and is hence the graph of some operator. We call it the **closure**
of \( T \), and denote it by \( \overline{T} \).

We are now in a position to define the **adjoint**. Let \( T \) be an operator
on \( H \), and notice that if we are successful in defining \( T^* \) so that \( \langle Tx , y \rangle = \langle x , T^*y \rangle \), for \( x \in D(T) \) and \( y \) somewhere, then the functional \( \phi(x) = \langle Tx , y \rangle \)
will be bounded, since \(|\phi(x)| \cdot ||x|| \cdot ||T^*y|| \). Thus we take as domain of the adjoint

\[ D(T^*) = \{ y \in K : x \rightarrow \langle Tx , y \rangle \text{ is a bounded linear functional on } D(T) \} \]

If \( D(T) \) is dense in \( H \), this extends to a bounded linear functional on \( H \), so by
the Riesz Representation theorem, there exists a unique \( x^* = T^*y \in H \) such
that \( \langle Tx , y \rangle = \langle x , T^*y \rangle \). By the uniqueness assertion, \( T^* \) is a linear
operator, and is called the **adjoint** of \( T \).

Stating that an operator is **self-adjoint** means two things. First,
\( D(T^*) \), defined above, is equal to \( D(T) \), and second, \( \langle Tx , y \rangle = \langle x , Ty \rangle \) for all
\( x \in D(T) \) and all \( y \in D(T^*) \).

We say that an operator \( T \) on \( H \) is **symmetric** if it satisfies the related
condition \( \langle Tx , y \rangle = \langle x , Ty \rangle \) for all \( x , y \in D(T) \). Please note that because of
the dependence on domain, there is a difference between a symmetric and a self-adjoint operator. This will become apparent in the examples which follow. However, we do have the following proposition.

**Proposition** Let $T$ be a densely defined operator. Then $T$ is symmetric $\iff T \subseteq T^\ast \iff \langle Tx, x \rangle \in \mathbb{R}$ for any $x \in \mathcal{D}(T)$.

**proof:** If $T$ is symmetric and $y \in \mathcal{D}(T)$, then the functional $\phi(x) = \langle Tx, y \rangle = \langle x, Ty \rangle$ satisfies $|\phi(x)| \leq \|x\| \|Ty\|$ and thus is bounded. This establishes that $\mathcal{D}(T) \subseteq \mathcal{D}(T^\ast)$. Since $\langle Tx, y \rangle = \langle x, Ty \rangle$ for $x, y \in \mathcal{D}(T)$ we see that $T \subseteq T^\ast$. It is also clear that if $T \subseteq T^\ast$, then $T$ is symmetric. Furthermore, if $T$ is symmetric, then for $x \in \mathcal{D}(T)$, $\langle Tx, x \rangle = \langle x, Tx \rangle = \langle Tx, x \rangle$, so $\langle Tx, x \rangle \in \mathbb{R}$. We leave as an exercise the argument showing that $\langle Tx, x \rangle \in \mathbb{R} \Rightarrow T$ is symmetric.

We'll see shortly that adjoints are always closed, so one consequence of this proposition is that every densely defined symmetric operator has a closed extension. A symmetric operator is said to be **maximally symmetric** if it has no proper symmetric extensions. It is not hard to see that a self-adjoint operator is maximally symmetric (just look at domains).

We now consider some examples, the first is a multiplication operator on $l_2$. Let $\{a_n\}$ be an unbounded sequence, let $\mathcal{D}(M) = \{ x \in l_2 : \sum a_n \langle x, e_n \rangle e_n \in l_2 \}$, and define $M$ by $Mx = \sum a_n \langle x, e_n \rangle e_n$. Notice that $e_n \in \mathcal{D}(M)$ for all $n$, so $\mathcal{D}(M)$ is dense. Furthermore, since $\{a_n\}$ is not bounded, $\mathcal{D}(M)$ is a proper subspace of $l_2$. Now fix $x, y \in \mathcal{D}(M)$. Then the functional $\phi(x) = \langle Mx, y \rangle = \sum a_n \langle x, e_n \rangle \langle e_n, y \rangle$, so $|\phi(x)| = |\langle Mx, y \rangle| = \sum a_n |\langle x, e_n \rangle| |\langle e_n, y \rangle| = \sum a_n |\langle x, e_n \rangle| |\langle e_n, y \rangle|$. Since $y$ was assumed to belong to $\mathcal{D}(M)$, this last term is finite, and $\phi$ is continuous. This implies that $\mathcal{D}(M) \subseteq \mathcal{D}(M^\ast)$. On the other hand, if $y \in \mathcal{D}(M^\ast)$, then there exists $K < \infty$ such that with $\phi(x) = \langle Mx, y \rangle = \sum a_n \langle x, e_n \rangle \langle e_n, y \rangle$, we have $|\phi(x)| = |\sum a_n \langle x, e_n \rangle \langle e_n, y \rangle| = |\langle x, \sum a_n \langle e_n, y \rangle e_n \rangle| \leq K\|x\|$. It follows that $\sum a_n \langle e_n, y \rangle e_n$ belongs to $l_2$, and in fact has norm no larger than $K$. But this shows that $\mathcal{D}(M^\ast) \subseteq \mathcal{D}(M)$. Next, the formula $\langle Mx, y \rangle = \langle x, M^\ast y \rangle$ implies that $M^\ast$ is simply multiplication
by the sequence \( \{ a_n \} \). Since \( \mathcal{D}(M) = \mathcal{D}(M^*) \), we see that \( M \) is self adjoint if and only if \( \{ a_n \} \subset \mathbb{R} \).

The next example is really three examples. We’ll consider three operators on \( L^2 [0,1] \), each defined by the formula \( T_k f = if' \), but differing in their domains. You’ll see that the properties of the operator depend very much on the domain. The domains of the three operators are as follows (\( AC[0,1] \) denotes the space of absolutely continuous functions):

\[
\mathcal{D}(T_1) = \{ f \in L^2 : f \in AC[0,1] \text{ and } f' \in L^2 \},
\]

\[
\mathcal{D}(T_2) = \{ f \in L^2 : f \in AC[0,1], f' \in L^2, \text{ and } f(0) = f(1) \},
\]

\[
\mathcal{D}(T_3) = \{ f \in L^2 : f \in AC[0,1], f' \in L^2, \text{ and } f(0) = f(1) = 0 \}.
\]

Let’s begin with a few observations. There are many ways to see that all three domains are dense, in fact, each contains a set of polynomials which is dense in \( L^2 \). It is evident that \( \mathcal{D}(T_3) \subset \mathcal{D}(T_2) \subset \mathcal{D}(T_1) \), so \( T_3 \subset T_2 \subset T_1 \). Next notice that \( \mathcal{R}(T_1) = L^2 \), while \( \mathcal{R}(T_2) = \mathcal{R}(T_3) = (\text{constants})^\perp \).

Next we see that

\[
\langle T_k f, g \rangle = \frac{1}{0} \int (if') \overline{g} \, dx
\]

\[
= i f \overline{g} \bigg|_0^1 - i \int f \overline{g} \, dx
\]

\[
= i f \overline{g} \bigg|_0^1 + \int f \overline{g'} \, dx.
\]
Thus whenever the boundary terms vanish we obtain a formula $\langle T_k f, g \rangle = \langle f, T_j g \rangle$. The boundary terms certainly disappear if (a) either $f$ or $g$ belongs to $\mathcal{D}(T_3)$, or (b) both $f$ and $g$ belong to $\mathcal{D}(T_2)$.

From the above discussion we see at least that $T_2$ is symmetric. We claim that in fact $T_1^* = T_3$, $T_2^* = T_2$, and $T_3^* = T_1$. Thus $T_2$ is self-adjoint, and $T_3$ is symmetric. In order to do this we need to check equality of domains, and take care of the boundary terms. The easiest domain inclusions are $\mathcal{D}(T_1) \subset \mathcal{D}(T_3^*)$, $\mathcal{D}(T_2) \subset \mathcal{D}(T_2^*)$, and $\mathcal{D}(T_3) \subset \mathcal{D}(T_1^*)$, so we'll do them first. If $g \in \mathcal{D}(T_3)$, then boundary terms disappear and $\langle T_1 f, g \rangle = \langle f, Tg \rangle$, so $\phi(f) = \langle T_1 f, g \rangle = \langle f, Tg \rangle$ satisfies $|\phi(f)| = \int_\mathcal{D}(T_3^*)$ and $T_3 \subset T_1^*$. If $f, g \in \mathcal{D}(T_2)$, then again $\phi(f) = \langle T_2 f, g \rangle = \langle f, Tg \rangle$ is bounded by $|\langle g, \phi(f) \rangle|$, so $g \in \mathcal{D}(T_2^*)$. Therefore $\mathcal{D}(T_2) \subset \mathcal{D}(T_2^*)$. The same argument works once more to show $\mathcal{D}(T_1) \subset \mathcal{D}(T_3^*)$.

For the reverse domain inclusions, assume $g \in \mathcal{D}(T_j^*)$, let $h = T_j^*g$, and let $H = \int_0^x h(t) \, dt$ be the anti-derivative of $h$. Now when $f \in \mathcal{D}(T_1)$,

$\langle T_1 f, g \rangle = \int i f' \overline{g} = \langle f, T_1^*g \rangle = \langle f, h \rangle$,

$= \int f \overline{h}$

$= f(1) \overline{H(1)} - f(0) \overline{H(0)} - \int f' \overline{(H)}$.

Now we may take $f$ to be a non-zero constant function (they're in $\mathcal{D}(T_1)$) and obtain $0 = f(1) \overline{H(1)} - 0$, so $H(1) = 0$. Therefore

$\int i f' \overline{g} = -\int f' \overline{H}$, so
\[ \int (i f') (g + i H) = 0. \]

Therefore \( g + iH \) is orthogonal to \( \mathcal{R}(T_1) \). But \( \mathcal{R}(T_1) \) is dense, so \( i g = H \).

But we have shown \( H(1) = 0 \), and from its definition, \( H(0) = 0 \). Thus \( H \), and with it, \( g \), belongs to \( \mathcal{D}(T_3) \). Thus at long last we’ve shown that \( T_1^* = T_3 \).

If \( f \in \mathcal{D}(T_2) \), and \( g \in \mathcal{D}(T_2^*) \), we proceed as above, again obtaining \( H(1) = H(0) = 0 \). This time, however, \( g + iH \in \mathcal{R}(T_2)^\perp \) implies only that

\[ g + iH \text{ is constant.} \]

Thus \( g = c \int h + D \). Then \( g(0) = D \), and \( g(1) = cH(1) + D \)

\[ = D, \text{ so } g \in \mathcal{D}(T_2). \]

Therefore \( \mathcal{D}(T_2^*) \subset \mathcal{D}(T_2) \), and \( T_2 = T_2^* \).

Finally, with \( f \in \mathcal{D}(T_3) \) and \( g \in \mathcal{D}(T_3^*) \), we again find that \( g + iH \) is constant. But it’s easy to belong to \( \mathcal{D}(T_1) \), and \( g = -iH + \text{constant does.} \)

This shows \( \mathcal{D}(T_3^*) \subset \mathcal{D}(T_1) \) and concludes this example.

Exercises

1. Some care about domains must be taken when using algebraic operations on unbounded operators. Prove that \( (T_1 T_2)T_3 = T_1(T_2T_3) \), and that if \( T_1 \subset T_2 \) then \( T_1 S \subset T_2 S \) and \( ST_1 \subset ST_2 \). Show that if \( S, T, \) and \( ST \) are densely defined, then \( T^* S^* \subset (ST)^* \).

2. Give an example of an operator \( T \) with graph \( G(T) \) dense in \( H \oplus H \). Discuss \( T^* \).

21. Symmetric and Self-adjoint operators

In this section we begin to prepare for the spectral theorem. We’ll discuss symmetric and self-adjoint operators, discuss invertibility, and define the resolvent and spectrum of an unbounded operator. We will make much use of the relation between the graph of \( T \) and the graph of \( T^* \) which is described in the next proposition.

**Proposition** Let \( T \) be a densely defined operator on \( H \), and define \( V: H \oplus H \rightarrow H \oplus H \) by \( V(x,y) = (-y,x) \). Then \( V \) is an isometry and \( G(T^*) = (VG(T))^\perp \).
proof: V is a rotation of $H \oplus H$, so it is an isometry. Since $\mathcal{D}(T)$ is dense, $T^*$ is defined, and the graphs of $T$ and $T^*$ are $G(T) = \{ (h,\text{Th}): h \in \mathcal{D}(T) \}$, and $G(T^*) = \{ (k,T^*k): k \in \mathcal{D}(T^*) \}$. Now if $h \in \mathcal{D}(T)$ and $k \in \mathcal{D}(T^*)$, then

$$\langle (k,T^*k), V(h,\text{Th}) \rangle = \langle (k,T^*k), (-\text{Th},h) \rangle$$

$$= -\langle k, \text{Th} \rangle + \langle T^*k, h \rangle$$

$$= -\langle k, \text{Th} \rangle + \langle k, \text{Th} \rangle = 0.$$  

Since $(h,\text{Th}) \in G(T)$, this implies that $(k,T^*k) \in [V \mathcal{G}(T)]^\perp$, so $G(T^*) \subset [V \mathcal{G}(T)]^\perp$. Conversely, if $(k,h) \in [V \mathcal{G}(T)]^\perp$, then for any $f \in \mathcal{D}(T)$, $0 = \langle (k,h), (-\text{Tf},f) \rangle = \langle k, -\text{Tf} \rangle + \langle h, f \rangle$, so $\langle \text{Tf}, k \rangle = \langle f, h \rangle$. Thus $\phi(f) = \langle \text{Tf}, k \rangle = \langle f, h \rangle$ is a bounded linear functional ($\| \phi(f) \| \cdot \| h \|$), so $k \in \mathcal{D}(T^*)$ and $T^*k = h$. Thus $(h,k) \in G(T^*)$ and we have $[V \mathcal{G}(T)]^\perp \subset G(T^*)$.

**Corollary** If $T$ is a densely defined operator on $H$, then $T^*$ is closed. In particular, self-adjoint operators are closed.

proof: Orthogonal complements are always closed, so $G(T^*)$ is closed.

**Corollary** If $T$ is a densely defined closed operator, then

$$H \oplus K = V \mathcal{G}(T) \oplus G(T^*).$$

In particular, the equations

$$- Tx + y = a$$

$$x + T^*y = b$$

are uniquely solvable for any $a, b \in H$. The solutions $x, y \in \mathcal{D}(T)$ depend linearly on $a$ and $b$.

proof: Since $T$ is closed, $G(T)$ is closed, and since $V$ is an isometry, $V \mathcal{G}(T)$ is closed. Thus $H \oplus K = V \mathcal{G}(T) \oplus G(T^*)$, and by this direct sum, for any $(a,b) \in H \oplus H$, there exist unique $x, y \in \mathcal{D}(T)$ with $(-Tx,x) + (y,Ty) = (a,b)$. But this is equivalent to the equations listed above.

We have seen that $T^*$ is defined whenever $\mathcal{D}(T)$ is dense, and the next result uses the rotation to give conditions for $T^*$ to be densely defined.
**Proposition** Let $T$ be a densely defined operator on $H$. Then

a. $T^*$ is densely defined $\iff$ $T$ is closeable.

b. If $T$ is closeable, then $\overline{T} = T^{**}$.

**proof:**

a. Suppose $T$ is closeable, and let $k \in \mathcal{D}(T^*)^\perp$. We wish, of course, to show that $k = 0$. If $x \in \mathcal{D}(T^*)$, then $\langle (k,0) , (x,T^*x) \rangle = \langle k , x \rangle + \langle 0 , T^*x \rangle = 0$. Thus $(k,0) \in G(T^*)^\perp = [\mathcal{V}G(T)]^\perp = \overline{\mathcal{V}G(T)} = \mathcal{V}(G(T))$. Since $T$ is closeable, $G(T)$ is a graph, and we see that $k = 0$.

Conversely, if $\mathcal{D}(T^*)$ is dense, then $T^{**}$ is defined and closed. It is not hard to see that $T \subset T^{**}$, so $T$ is closeable.

b. We need only check that $G(T^{**}) = \overline{G(T)}$. Now $G(T^{**}) = [\mathcal{V}[G(T^*)]^\perp]^\perp = [\mathcal{V}[G(T)]^\perp]^\perp = \mathcal{V}[G(T)]^\perp = \mathcal{V}(G(T)) = \overline{G(T)}$.

**Corollary** If $T$ is closed, then $T^*$ is densely defined and $T = T^{**}$.

Just as in the case of bounded operators, the range of an operator is related to the kernel of its adjoint.

**Proposition** If $T$ is densely defined, then $\mathcal{R}(T)^\perp = \text{Ker}T^*$. If $T$ is closed, then $\mathcal{R}(T^*)^\perp = \text{Ker}T$.

**proof:**

a. If $h \in \text{Ker}T^*$, then whenever $x \in \mathcal{D}(T)$, $0 = \langle x , T^*h \rangle = \langle Tx , h \rangle$, so $h \in \mathcal{R}(T)^\perp$. On the other hand, if $h \in \mathcal{R}(T)^\perp$, then for any $x \in \mathcal{D}(T)$, $\langle Tx , h \rangle = 0$, so $x \rightarrow \langle Tx , h \rangle$ is about as bounded as a functional can be. Thus $h \in \mathcal{D}(T^*)$ and $T^*h = 0$.

b. If $T$ is closed, then $T^{**} = T$, and by part a, $\mathcal{R}(T^*)^\perp = \text{Ker}T^{**} = \text{Ker}T$.

If $T$ is an operator on $H$, we say that $T$ is **boundedly invertible**, or simply **invertible**, if there is a bounded linear operator $S$ (usually denoted by $T^{-1}$) such that $TS = I$, and $ST \subset I$. It is not hard to see that an operator $T$ can have at most one bounded inverse. Furthermore,

**Proposition** An operator $T$ is boundedly invertible if and only if

a. $\text{Ker}T = \{0\}$,

b. $\mathcal{R}(T) = H$, and
c. $G(T)$ is closed.

proof: Suppose $S$ is the bounded inverse of $T$. Since $ST \subset I$, we have that $\ker T = \{0\}$. Since $TS = I$, $\mathcal{R}(T) = H$, and $G(T) = \{ (x, Tx) : x \in D(T) \} = \{ (Sy, y) : y \in H \}$, which is closed since $S$ is bounded. In the other direction, by a. and b., $T^{-1}$ is well-defined, and since $G(T^{-1})$ is closed, $T^{-1}$ is bounded by the closed graph theorem.

The **resolvent set** of a linear operator on $H$ is defined to be the set of all $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is boundedly invertible. The **spectrum** of $T$, still denoted by $\sigma(T)$, is the complement of the resolvent set. You should note that if $T$ is not closed, then $T - \lambda I$ isn’t either, so $\sigma(T) = \mathbb{C}$. Thus the spectrum of an operator need not be a bounded set. It is also possible that $\sigma(T) = \emptyset$, but $\sigma(T)$ is always a closed subset of $\mathbb{C}$. We invite you to compute the spectra of the operators given in the examples in Section 20.

**Proposition** If $T$ is a symmetric operator on $H$, then

a. If $\mathcal{R}(T)$ is dense, then $T$ is 1-1.

b. If $T$ is 1-1 and self-adjoint, then $\mathcal{R}(T)$ is dense and $T^{-1}$ is self-adjoint. ($T^{-1}$ is not necessarily bounded.)

c. If $D(T) = H$, then $T$ is self-adjoint and bounded.

d. If $\mathcal{R}(T) = H$, then $T$ is self-adjoint and $T^{-1}$ is bounded.

proof: a. Since $\ker T^* = \mathcal{R}(T)^\perp = \{0\}$, $T^*$ is 1-1. Since $T$ is symmetric, $T^*$ is an extension of $T$, so $T$ must be 1-1 also.

b. $T^* = T$ is 1-1, so $\mathcal{R}(T)^\perp = \ker T^* = \{0\}$, and hence $\mathcal{R}(T)$ is dense in $H$. Thus $D(T^{-1}) = \mathcal{R}(T)$ is dense and $T^{-1}$ is defined. Since $T$ is self-adjoint we have that $\langle T^{-1} x, y \rangle = \langle x, T^{-1} y \rangle$. Notice that for any operator $S$, $VG(-S) = \{ (Sx, x) \} = G(S^{-1})$. Similarly, $VG(S^{-1}) = G(-S)$. Now since $T$ is self-adjoint, $T$ is closed, and so is $-T$. Therefore

$$H \oplus H = G(-T^*) \oplus VG(-T) = G(-T) \oplus G(T^{-1}).$$

But

$$H \oplus H = G(T^{-1})^* \oplus VG(T^{-1}) = G(T^{-1})^* \oplus G(-T).$$

Thus $G(-T)^\perp = G(T^{-1}) = G(T^{-1})^*$, so $T^{-1}$ is self-adjoint.
c. Since $T$ is symmetric, $T \subset T^*$, but since $\mathcal{D}(T) = H$, $\mathcal{D}(T) = \mathcal{D}(T^*)$, so $T$ is self-adjoint. Thus $T$ is closed, so again since $\mathcal{D}(T) = H$, $T$ is continuous by the closed graph theorem.

d. By part a, $T$ is 1-1, and $T^{-1}$ is well defined, with $\mathcal{D}(T^{-1}) = H$. Since $T^{-1}$ is symmetric, it is self-adjoint and bounded by c.

The next lemma will be useful in a couple of contexts, first to show that the spectrum of a self-adjoint operator is real, and later to define the Cayley transform of a self-adjoint operator.

**Proposition** Let $T$ be a symmetric operator, and let $\lambda = \alpha + i \beta$. Then

a. $\forall x \in \mathcal{D}(T)$, $\| (T - \lambda I)x \|^2 = \| (T - \alpha I)x \|^2 + |\beta|^2 \| x \|^2$.

b. If $\beta \neq 0$, then $T - \lambda I$ is 1-1.

c. If $\beta = 0$ and $T$ is closed, then $\mathcal{R}(T - \lambda I)$ is closed.

**proof:** We have

$$
\| (T - \alpha I)x - i \beta x \|^2 = \langle (T - \alpha I)x - i \beta x, (T - \alpha I)x - i \beta x \rangle
$$

$$
= \| (T - \alpha I)x \|^2 + |\beta|^2 \| x \|^2 + \langle (T - \alpha I)x, i \beta x \rangle + \langle -i \beta x, (T - \alpha I)x \rangle
$$

$$
= \| (T - \alpha I)x \|^2 + |\beta|^2 \| x \|^2 + 2 \text{Re} i \langle (T - \alpha I)x, \beta x \rangle.
$$

Since $T$ is symmetric, this last inner product is real, and we obtain

$$
\| (T - \alpha I)x - i \beta x \|^2 = \| (T - \alpha I)x \|^2 + |\beta|^2 \| x \|^2.
$$

Now this implies that

$$
\| (T - \alpha I)x - i \beta x \|^2 \cdot |\beta|^2 \| x \|^2,
$$

so if $\beta \neq 0$, ker($T - \lambda I$) = {0}, and $T - \lambda I$ is 1-1. To prove c., notice that the norm on $G(T - \alpha I)$ is $\| (T - \alpha I)x \|^2 = \| (T - \alpha I)x \|^2 + \| x \|^2$, and hence

$$
\min(1, |\beta|) (\| (T - \alpha I)x \|^2 + \| x \|^2) \cdot
$$

$$
\| (T - \lambda I)x \|^2 \cdot \max(1, |\beta|) (\| (T - \alpha I)x \|^2 + \| x \|^2).
$$
Since \( \| (T - \lambda I) x \|^2 \) is the norm on \( \mathcal{R}(T - \lambda I) \), the above inequalities show that for \( \beta \cdot 0, \mathcal{R}(T - \lambda I) \) is closed if and only if \( G(T - \alpha I) \) is closed. But since \( G(T) \) is assumed to be closed, \( G(T - \alpha I) \) is closed.

We wish to point out some consequences of this last result. If \( T \) is self-adjoint, then \( T \) is both symmetric and closed, and as a result we see that (taking \( \lambda = \pm i \) ) \( \ker(T \pm i I) = \{0\} \) and \( \mathcal{R}(T \pm i I) \) is closed. We shall make extensive use of this in the next Section. Another consequence is that the spectrum of a self-adjoint operator is real.

**Proposition** If \( T \) is a self-adjoint operator, then \( \sigma(T) \subset \mathbb{R} \).

**proof:** Suppose \( \lambda = \alpha + i \beta \), with \( \beta \cdot 0 \). Then \( T - \lambda I \) is 1-1, and has closed range. Furthermore, \( \mathcal{R}(T - \lambda I)^\perp = \ker(T - \overline{\lambda} I) = \{0\} \), so \( \mathcal{R}(T - \lambda I) \) is dense. We thus have that \( T - \lambda I \) is 1-1, \( \mathcal{R}(T - \lambda I) = \mathcal{H} \), and \( G(T - \lambda I) \) is closed. It follows that \( T - \lambda I \) is boundedly invertible, and \( \lambda \) is in the resolvent set of \( T \).

More can be proved along these same lines. We leave the details to the reader, but one can show that \( \dim(\ker(T^* - \lambda I)) \) is constant on the two half planes \( \mathbb{C}^u = \{ z \in \mathbb{C}: \text{Im} z > 0 \} \) and \( \mathbb{C}^l = \{ z \in \mathbb{C}: \text{Im} z < 0 \} \), and from this one deduces that if \( T \) is a closed, symmetric operator, then \( \sigma(T) \) is either \( \mathbb{C} \), the closure of \( \mathbb{C}^u \), the closure of \( \mathbb{C}^l \), or a subset of \( \mathbb{R} \). Furthermore, one can obtain a converse to our last proposition: If \( T \) is closed and symmetric, then \( T \) is self-adjoint \( \iff \sigma(T) \subset \mathbb{R} \), \( \iff \ker(T^* - iI) = \ker(T^* + iI) = \{0\} \).

**Exercises**

1. Prove that the spectrum of an operator is always a closed subset of \( \mathbb{C} \).
   **Suggestion:** Assume \( \lambda_0 \) is in the resolvent set, and then try to expand \( (\lambda I - T)^{-1} \) in a series in \( (\lambda - \lambda_0) \) for \( \lambda \) near \( \lambda_0 \).

2. Prove that any closed subset of \( \mathbb{C} \) is the spectrum of some operator.

3. Let \( T_1, T_2, T_3 \) be as in the examples, and let \( T_4 \) be defined by \( T_4 f = i f' \) with \( \mathcal{D}(T_4) = \{ f : f \in AC[0,1], f' \in L^2, f(0) = 0 \} \). What are \( \sigma(T_k) \) and \( \sigma_p(T_k) \) for \( k = 1, 2, 3, 4 \)?
4. Prove that if $T$ is a symmetric operator on $H$ there is a maximal symmetric operator $\hat{T}$ which extends $T$.

22. The Cayley Transform

We will obtain the spectral theorem for unbounded self-adjoint operators from the theory for (bounded) unitary operators. The function $f(\lambda) = \frac{\lambda - i}{\lambda + i}$ maps the real axis onto the unit circle $\mathbb{T}$ with the point $\{1\}$ removed. Since the spectrum of a self-adjoint operator is real, we expect that if we can define $f(T)$, then it will be a unitary operator. We’ll then carry the spectral resolution for $f(T)$ back to $T$.

In order to carry out this program, we’ll begin by defining the Cayley Transform of a self-adjoint operator and showing how to integrate measurable (but unbounded) functions with respect to a resolution of the identity. This will keep us out of trouble for the next two sections, and then we’ll prove the spectral theorem for self-adjoint operators.

To begin with we state a special case of a result from the last section.

**Proposition** Let $T$ be a symmetric operator on $H$. Then

a. $\| (T \pm iI)x \|^2 = \|Tx\|^2 + \|x\|^2 \quad \forall x \in \mathcal{D}(T)$.

b. $T \pm iI$ is 1-1.

c. The map $(T \pm iI)x \rightarrow (x, Tx)$ is an isometry from $\mathcal{R}(T \pm iI)$ onto $\mathcal{G}(T)$. Thus $\mathcal{R}(T \pm iI)$ is closed if and only if $T$ is closed.

We will also need the special case $T = T^*$ of the next theorem.

**Theorem** If $T$ is a densely defined closed operator on $H$, then the operators $B = (I + T^*T)^{-1}$ and $C = T(I + T^*T)^{-1}$ are defined on all of $H$, and $\|B\| \cdot 1$, $\|C\| \cdot 1$. $B$ is a self-adjoint positive operator.

**proof:** Since $H \oplus H = G(T^*) \oplus VG(T)$, as we noted earlier we can solve the equations

\[ -Tf + g = 0, \]
\[ f + T^*g = h. \]
for any $h \in H$. We define the linear operators $B$ and $C$ by $f = Bh$ and $g = Ch$, and by means of the orthogonal decomposition $H \oplus H = G(T^*) \oplus VG(T)$, see that

$$\| h \|^2 = \| (h,0) \|^2 = \| (f,Tf) \|^2 + \| (-T^*g,g) \|^2 = \| f \|^2 + \| Tf \|^2 + \| g \|^2 + \| T^*g \|^2$$

Thus

$$\| Bh \|^2 + \| Ch \|^2 = \| f \|^2 + \| g \|^2 \cdot \| h \|^2,$$

so $B$ and $C$ are both of norm no larger than 1. For any $x \in \mathcal{D}(T^*T)$,

$$\langle (I + T^*T)x, x \rangle = \langle x, x \rangle + \langle Tx, Tx \rangle \cdot \langle x, x \rangle,$$

so $I + T^*T$ is 1-1. and thus $(I + T^*T)^{-1}$ is defined. The equations above may be rewritten as

$$-TB + C = 0,$$
$$B + T^*C = I,$$

so that

$$C = TB,$$
$$I = B + T^*TB = (I + T^*T)B.$$

The second of these imply that $B$ is everywhere defined and is $(I + T^*T)^{-1}$, and the first then shows that $C = T(I + T^*T)^{-1}$. We leave the other assertions regarding $B$ to the reader.

**Proposition** Suppose $U$ is an operator on $H$ which satisfies $\| Ux \| = \| x \|$ for any $x \in \mathcal{D}(U)$. Then

a. $x, y \in \mathcal{D}(U) \Rightarrow \langle Ux, Uy \rangle = \langle x, y \rangle$.

b. $\mathcal{R}(I - U)$ dense $\Rightarrow I - U$ is 1-1.

c. If any one of $\mathcal{D}(U), \mathcal{R}(U), G(U)$ is closed, so are the others.
proof: Part a is simply the polarization identity. As for b., if \( x \in \mathcal{D}(U) \) and 
\[(I - U)x = 0, \text{ then } x = Ux, \text{ and } \langle x, (I-U)x \rangle = \langle x, x \rangle - \langle x, Uy \rangle = \langle Ux, Uy \rangle - \langle x, Uy \rangle = 0. \text{ Thus } x \in \mathbb{R}(I - U)^\perp = \{0\}, \text{ so } U \text{ is 1-1.} \]
Since \( U \) is an isometry on its domain, \( \mathcal{D}(U) \) is closed if and only if \( \mathbb{R}(U) \) is closed. If \( \mathcal{D}(U) \) is closed, then \( U \) is bounded so its graph is closed. We leave the remaining implication to the reader.

Now to the definition of the Cayley transform. Since
\[
|| (T + iI)x ||^2 = || (T - iI)x ||^2,
\]
we may define an isometry \( U: \mathbb{R}(T + iI) \to \mathbb{R}(T - iI) \) by
\[
U((T + iI)x) = (T - iI)x.
\]
Then \( T + iI \) maps \( \mathcal{D}(T) \) in a 1-1 fashion to \( \mathbb{R}(T + iI) = \mathcal{D}(U) \), so we may define
\[
(T + iI)^{-1} : \mathcal{D}(U) \to \mathcal{D}(T)
\]
and obtain \( U = (T - iI)(T + iI)^{-1} \) which maps \( \mathcal{D}(U) = \mathbb{R}(T + iI) \to \mathbb{R}(U) = \mathbb{R}(T - iI) \). \( U \) is called the Cayley transform of \( T \).

**Theorem** Let \( T \) be a symmetric operator (we need not assume \( T \) is densely defined), and let \( U = (T - iI)(T + iI)^{-1} \) be its Cayley transform. Then
\begin{itemize}
  \item[a.] \( U \) is closed \( \iff \) \( T \) is closed.
  \item[b.] \( \mathbb{R}(I - U) = \mathcal{D}(T) \),
  \( I - U \) is 1-1, and
  \[ T = i(I + U)(I - U)^{-1}. \]
  \item[c.] \( U \) is unitary \( \iff \) \( T \) is self-adjoint.
\end{itemize}
Conversely, if \( W \) is isometric on its domain and if \( (I - W) \) is 1-1, then \( W \) is the Cayley transform of a symmetric operator on \( H \).

**proof:** a. We know that \( T \) is closed \( \iff \) \( \mathbb{R}(T + iI) \) is closed, and \( U \) is closed \( \iff \) \( \mathcal{D}(U) \) is closed. Since \( \mathcal{D}(U) = \mathbb{R}(T + iI) \), a is proved.

b. \( T + iI \) provides a 1-1 correspondence between \( \mathcal{D}(T) \) and \( \mathcal{D}(U) \).
Specifically
\[
z = (T + iI)x \quad \text{ and } \quad Uz = (T - iI)x.
\]
Adding and subtracting these yields
\[(I - U)z = 2ix \quad \text{and} \quad (I + U)z = 2Tx.\]

Thus \(\mathcal{R}(I - U) = \mathcal{D}(T)\), and \(I - U\) is 1-1. Furthermore, \((I - U): \mathcal{D}(U) \to \mathcal{R}(I - U) = \mathcal{D}(T)\). Thus \((I - U)^{-1}: \mathcal{D}(T) \to \mathcal{D}(U)\) and \(Tx = i(I + U)(I - U)^{-1} x\) for all \(x \in \mathcal{D}(T)\). Thus \(b\) is proved.

c. Suppose \(U\) is unitary. We will first show that in this case \(T\) is densely defined, so \(T^*\) is in fact defined. Notice that \(I - U\) is normal, so \(\ker(I - U) = \ker(I - U^*)\). We have \((\mathcal{R}(I - U)) \perp = \ker(I - U^*) = \ker(I - U) = \{0\}\) by part \(b\)., so \(\mathcal{R}(I - U) = \mathcal{D}(T)\) is dense. Since \(T\) is symmetric, \(T \subset T^*\), so to show \(T\) is self-adjoint, we need only show that \(T^* \subset T\), and this is a question of domains. Since \(U\) is unitary, \(U\) is closed, and thus by part \(a\), \(T\) is closed. Hence \(\mathcal{R}(T + iI)\) is closed. Furthermore, \(\mathcal{R}(T + iI)\) is dense, for if \(x \in \mathcal{R}(T + iI)^\perp\), then \(0 = \langle (T + iI)y, x \rangle = \langle y, (T - iI)x \rangle\) for all \(y \in \mathcal{D}(T)\). Since \(\mathcal{D}(T)\) is dense, \(x \in \ker(T - iI)\), which is \(\{0\}\) since \(T - iI\) is 1-1. It follows that if \(U\) is unitary, then \(\mathcal{R}(T + iI) = \mathcal{H}\). Now let \(y \in \mathcal{D}(T^*)\). Since \(T + iI\) is onto \(\mathcal{H}\), there exists \(y_0\) with

\[
(T + iI) y_0 = (T^* + iI) y
\]

\[
= (T^* + iI) y_0 \quad \text{(since \(T \subset T^*\)).}
\]

Setting \(y_1 = y - y_0\), \(y_1 \in \mathcal{D}(T^*)\), and

\[
( (T - iI) x , y_1 ) = \langle x , (T^* + iI) y_1 \rangle = \langle x , 0 \rangle = 0.
\]

Thus \(y_1 \in \mathcal{R}(T - iI)^\perp = \{0\}\), so \(y = y_0 \in \mathcal{D}(T)\).

Now suppose \(T = T^*\) is self-adjoint. We have seen above then that \(I + T^2\) is 1-1, maps \(\mathcal{D}(T)\) onto \(\mathcal{H}\), and has a bounded inverse. But

\[
I + T^2 = (T - iI)(T + iI), \quad \text{and} \quad I + T^2 = (T + iI)(T - iI).
\]

The first of these shows that \(\mathcal{R}(U) = \mathcal{R}(T - iI) = \mathcal{R}(I + T^2) = \mathcal{H}\), while the second shows that \(\mathcal{D}(U) = \mathcal{R}(T + iI) = \mathcal{R}(I + T^2) = \mathcal{H}\). Since \(U\) preserves norms, this implies \(U\) is unitary.
Conversely, suppose $\|Wx\| = \|x\|$ for all $x \in \mathcal{D}(W)$ and that $I - W$ is 1-1. This last assumption of course means that 1 is not an eigenvalue of $W$, so the formula $x = z - Wz$ is a 1-1 correspondence between $\mathcal{D}(W)$ and $\mathcal{R}(I - W)$. Then each $x \in \mathcal{R}(I - W)$ is of the form $z - Wz$ for some $z$. In this case, define $T$ on $\mathcal{R}(I - W)$ by $Tx = i(z + Wz)$. Then for $x, y \in \mathcal{D}(T), x = z - Wz, y = u - Wu$, so

\[
\langle Tx, y \rangle = i \langle z + Wz, u - Wu \rangle \\
= i(\langle z, u \rangle + \langle Wz, u \rangle - \langle z, Wu \rangle - \langle Wz, Wu \rangle) \\
= i(\langle Wz, u \rangle - \langle z, Wu \rangle) \text{ (since $W$ preserves inner products),}
\]

and by the same calculations

\[
\langle x, Ty \rangle = \langle z - Wz, u + Wu \rangle \\
= i(\langle Wz, u \rangle - \langle z, Wu \rangle).
\]

Thus $\langle Tx, y \rangle = \langle x, Ty \rangle$, so $T$ is symmetric. It is easily checked that $(T - iI)x = W(T + iI)x$ for all $x \in \mathcal{D}(T)$ and $\mathcal{D}(W) = \mathcal{R}(T + iI)$, so that $W$ is the Cayley transform of $T$.

**Exercises**

1. Let $M$ be a multiplication operator, say by the sequence $\{a_n\} \subset \mathbb{R}$, on $l_2(\mathbb{Z})$. Compute the Cayley transform of $M$ and verify directly that it is unitary.

2. Let $M$ be the multiplication operator by the real-valued function $\phi$ on $L^2[0,1]$. Compute the Cayley transform of $M$ and verify directly that it is unitary.

2. Let $T$ be closed, symmetric, and densely defined. The **deficiency indices** of $T$ are the dimensions of $\mathcal{R}(T + iI)\perp$ and $\mathcal{R}(T - iI)\perp$. Show that $T$ is self-adjoint if and only if both deficiency indices are 0, that $T$ is maximally symmetric if and only if at least one index is zero, and that $T$ has a self-adjoint extension if and only if its deficiency indices are equal.

3. The right shift $S$ on $l_2(\mathbb{N})$ is an isometry and $I - S$ is 1-1, so $S$ is the Cayley transform of a symmetric operator $T$. What are the deficiency indices of $T$? Say as much as you can about $T$. 

74

School of Mathematics

Georgia Institute of Technology
23. Integration of (unbounded) measurable functions

If E is a resolution of the identity for \((X, \Omega, H)\), we have seen how to construct a representation \(\rho : \mathcal{B}(X) \to \mathcal{B}(H)\) from the bounded measurable functions on X into the bounded operators on H. The representation satisfies the formula 
\[
\langle \rho(f)x, y \rangle = \int f(\lambda) dE_{x,y}(\lambda)
\]
for each \(x, y \in H\). We now extend this representation to the space of measurable functions on X. Of course if \(f\) is not bounded, \(\rho(f)\) won’t be either. Throughout this section, when \(f\) is a measurable function, we let \(\omega_n = \{\lambda : |f(\lambda)| < n\}\), \(\chi_n = \chi_{\omega_n}\), and let \(f_n\) be the truncation \(f_n = f\chi_{\omega_n}\). We first define what will be the domain of \(\rho(f)\).

**Theorem** Let E be a resolution of the identity on \((X, \Omega, H)\), and let \(f\) be a complex-valued measurable function on X. Let

\[
\mathcal{D}_f = \{x \in H : \int_X |f|^2 dE_{x,x} < \cdot \}.
\]

Then

a. \(\mathcal{D}_f\) is a dense subspace of H.

b. \(\forall x, y \in H, \int_X |f| dE_{x,y} \cdot \|y\| \left[ \int_X |f|^2 dE_{x,x} \right]^{1/2}\).

c. If \(f\) is bounded and \(v = \rho(f)z\), then \(dE_{x,v} = f dE_{x,z}\) \(\forall x, z \in H\).

**proof:** In part b, the notation \(\|\mu\|\) is for the variation measure of \(\mu\). In part c, \(\rho\) refers to the representation of \(\mathcal{B}(X)\) constructed earlier, and will be used to prove the multiplicativity of the representation \(\rho\) we seek.

In the case when \(f\) is a characteristic function, we have \(E_{\omega,\omega}(\omega) = \langle E(\omega)(\omega), \omega \rangle = |\omega|^2 \cdot E(\omega)x, x\rangle\), so we see that \(\mathcal{D}_f\) is closed under scalar multiplication. As for closure under addition, since \(2ab \cdot a^2 + b^2\) for all real \(a, b\), we have

\[
|| E(\omega)(x + y) ||^2 \cdot (|| E(\omega)x || + || E(\omega)y ||)^2
\]

\[
\cdot 2 (|| E(\omega)x ||^2 + || E(\omega)y ||^2 )
\]
so that $D_f$ is closed under addition.

To see that $D_f$ is dense in $H$, let $\omega_n = \{\lambda: |f(\lambda)| < n\}$, let $f_n$ be the truncation $f_n = f \chi_{\omega_n}$. Since $f_n$ is bounded, $D_f \supseteq H$. Now suppose $x = E(\omega_n)x \in \mathcal{R}(E(\omega_n))$. Then $E(\omega)x = E(\omega)E(\omega_n)x = E(\omega \cap \omega_n)x$, so $E_{x,x}(\omega) = E_{x,x}(\omega \cap \omega_n)$ for each $\omega \in \Omega$. But this implies that

$$\int_X |f|^2 \, dE_{x,x} = \int_{\omega_n} |f|^2 \, dE_{x,x} \cdot n^2 \|x\|^2.$$ 

Thus for every $n$, $\mathcal{R}(E(\omega_n)) \subset D_f$. Since $\cup \omega_n = X$, the countable additivity of $E$ implies that $E(\omega)z = \lim E(\omega_n)z$ $\forall z \in H$, so $D_f$ is dense in $H$.

We will first prove b. in the case that $f$ is bounded. For any $x, y \in H$, the measure $f \, dE_{x,y}$ is absolutely continuous with respect to $|f| \, d|E_{x,y}|$, so there exists a function $u$ (which in this case has modulus 1) with $uf \, dE_{x,y} = |f| \, d|E_{x,y}|$. Therefore

$$\int_X |f| \, dE_{x,y} = \int u \, f \, dE_{x,y}$$

$$= \langle \rho(uf)x, y \rangle \quad (\text{since } f \text{ is bounded})$$

$$\cdot \| \rho(uf)x \| \| y \|$$

$$= \left[ \int |uf|^2 \, dE_{x,x} \right]^{1/2} \| y \|$$

$$= \left[ \int |f|^2 \, dE_{x,x} \right]^{1/2} \| y \|.$$ 

But this is b., in the case when $f$ is bounded. With $\omega_n, \chi_n$, and $f_n$ as before, we have

$$\int_X |f_n| \, dE_{x,y} \cdot \left[ \int |f_n|^2 \, dE_{x,x} \right]^{1/2} \| y \|.$$
Sending n → *, we obtain b.

As for c, let v = ρ(f), let g be bounded and measurable, and observe that

\[ \int g \, dE_{x,v} = \langle \rho(g)x , v \rangle = \langle \rho(g)x , \rho(f)z \rangle = \langle \rho(\bar{f}g)x , z \rangle = \int g \, \bar{f} \, dE_{x,z}. \]

Since g was an arbitrary bounded measurable function the measures dE_{x,v} and \( \bar{f} \, dE_{x,z} \) must be equal.

The next theorem is the main result of this section. Part d is referred to as the multiplication theorem. Part e includes a strengthening of the multiplication theorem for a function f and its conjugate \( \bar{f} \).

**Theorem** Let E be a resolution of the identity on \((X,\Omega,H)\). The representation \( \rho:B(X) \to B(H) \) may be extended to a representation defined on the measurable functions on X. For each measurable function f defined on X

a. \( \rho(f) \) is a densely defined operator with \( D(\rho(f)) = D_f \).

b. \( \langle \rho(f)x , y \rangle = \int_X f \, dE_{x,y} \quad \forall x \in D_f , \forall y \in H. \)

c. \( \| \rho(f)x \|^2 = \int_X |f|^2 \, dE_{x,x} \quad \forall x \in D_f. \)

d. For every f, g measurable, \( \rho(f) \rho(g) \subset \rho(fg) \) and \( D(\rho(f)\rho(g)) = D_g \cap D_{fg} \).

e. For any measurable f, \( \rho(f)^* = \rho(\bar{f}) \), so \( \rho(f) \) is closed. Furthermore, \( \rho(f)\rho(f)^* = \rho(|f|^2) = \rho(f)^*\rho(f) \).

**proof:** Parts a, b, and c come quickly; we'll have to work for the others. If \( x \in D_f \), then \( y \to \int f \, dE_{x,y} \) is a conjugate linear functional. It is bounded,

\[ \int_X |f| \, dE_{x,y} \cdot \int_X |f| \, dE_{x,y} \cdot \| y \| \left[ \int_X |f|^2 \, dE_{x,x} \right]^{1/2}. \]

This last

School of Mathematics
Georgia Institute of Technology
term is finite since \( x \in \mathcal{D}_f \). Thus there exists a unique element of \( H \), which we denote by \( \rho(f)x \), such that

\[
\langle \rho(f)x, y \rangle = \int f dE_{x,y} \quad \forall \ y \in H.
\]

The linearity of \( \rho(f) \) follows from the uniqueness of \( \rho(f)x \). We have now established a, and b. As for c, \( \forall \ x \in \mathcal{D}_f \),

\[
\| \rho(f)x \| = \sup \{ \| \langle \rho(f)x, y \rangle \| : \| y \| = 1 \}
\]

\[
= \left( \int f^2 dE_{x,x} \right)^{1/2} \cdot \left( \int \| f \|^2 dE_{x,x} \right)^{1/2}.
\]

To obtain the inequality in the other direction, since \( f_n \) is bounded, \( \mathcal{D}_f f_n = \mathcal{D}_f \) so \( \| \rho(f)x - \rho(f_n)x \|^2 \leq \int |f - f_n|^2 dE_{x,x} \). By countable additivity, this integral tends to 0 as \( n \) tends to \( \cdot \), so \( \rho(f_n)x \to \rho(f)x \), and as a consequence, \( \| \rho(f_n)x \| \to \| \rho(f)x \| \). Since \( f_n \) is bounded, \( \| \rho(f_n)x \|^2 = \int |f_n|^2 dE_{x,x} \). Thus \( \| \rho(f)x \|^2 = \int |f|^2 dE_{x,x} = \lim \int |f_n|^2 dE_{x,x} \cdot \int |f|^2 dE_{x,x} \) by Fatou’s Lemma. This completes the proof of part c.

As for the multiplication theorem, note first that \( \mathcal{D}(\rho(f)\rho(g)) = \{ x: x \in \mathcal{D}_g \text{ and } \rho(g) \in \mathcal{D}_f \} \). We now establish the multiplication theorem in steps. First, assume that \( g \) is measurable and \( f \) is both measurable and bounded. We then have \( \int |fg|^2 dE_{x,x} \cdot \| f \| \cdot \int |g|^2 dE_{x,x} \), so \( \mathcal{D}_{fg} \subset \mathcal{D}_g \) when \( f \) is bounded. In particular, \( \mathcal{D}_{fg} \cap \mathcal{D}_g = \mathcal{D}(g) = \mathcal{D}(\rho(f)\rho(g)) \). Now let \( z \in H \), take \( v = \rho(\overline{f})z \), and observe that

\[
\langle \rho(f)\rho(g)x, z \rangle = \langle \rho(g)x, \rho(\overline{f})z \rangle \quad \text{(since } f \text{ is bounded)}
\]

\[
= \int g dE_{x,v}
\]

\[
= \int fg dE_{x,z}
\]
Thus for \( f \) bounded and \( g \in \mathcal{D}_g \), we have \( \rho(f)\rho(g)x = \rho(fg)x \). Still operating under the assumption that \( f \) is bounded, let \( y = \rho(g)x \). Then

\[
\int |f|^2 dE_{y,y} = \| \rho(f)y \|^2
\]

\[
= \langle \rho(f)\rho(g)x, \rho(f)\rho(g)x \rangle
\]

\[
= \langle \rho(fg)x, \rho(fg)x \rangle \quad (\text{since } f \text{ is bounded})
\]

\[
= \| \rho(fg)x \|^2
\]

\[
= \int |fg|^2 dE_{x,x}.
\]

Applying the monotone convergence theorem, we obtain

\[
\int |f|^2 dE_{y,y} = \int |fg|^2 dE_{x,x} \quad \text{for all bounded measurable } f, g.
\]

Now \( \mathcal{D}(\rho(f)\rho(g)) = \{ x : x \in \mathcal{D}_g \text{ and } \rho(g) \in \mathcal{D}_f \} \), so \( y = \rho(g)x \in \mathcal{D}_f \iff x \in \mathcal{D}_{fg} \), so \( \mathcal{D}(\rho(f)\rho(g)) = \mathcal{D}_g \cap \mathcal{D}_{fg} \) as advertised. Continuing to let \( y = \rho(g)x \), \( f_n \to f \) in \( L^2(E_{y,y}) \) by the Lebesgue convergence theorem, so we have both

\[
\| \rho(f_n)y - \rho(f)y \|^2 \quad \cdot \quad \int |f_n - f|^2 dE_{y,y}, \text{ and}
\]

\[
\| \rho(f_n g)x - \rho(f_n)gx \|^2 \quad \cdot \quad \int |f_n g - fg|^2 dE_{x,x}.
\]

Then

\[
\rho(f)\rho(g)x \quad = \quad \rho(f)y \quad = \quad \lim \rho(f_n)y
\]

\[
= \lim \rho(f_n)\rho(g)x
\]

\[
= \lim \rho(f_n g)x
\]
\[ = \rho(fg)x, \]

so the multiplication theorem is proven.

Next, we will show that \( \mathcal{D}_{\overline{f}} = \mathcal{D}_f \subset \mathcal{D}(\rho(f)^* \rho(\overline{f})) \) and \( \mathcal{D}(\rho(f)) \subset \mathcal{D}(\rho(f)^*). \)

Indeed, if \( x \in \mathcal{D}_f \) and \( y \in \mathcal{D}_{\overline{f}} \), then

\[
\langle \rho(f)x, y \rangle = \lim \langle \rho(f_n)x, y \rangle \\
= \lim \langle x, \rho(f_n)y \rangle \quad \text{(since } f_n \text{ is bounded)} \\
= \langle x, \rho(f_\overline{n})y \rangle.
\]

Thus \( x \to \langle \rho(f)x, y \rangle \) is bounded and \( y \in \mathcal{D}(\rho(f)^*). \) The above calculation also shows that \( \rho(f) \subset \rho(f)^*. \) We must now show that \( \mathcal{D}(\rho(f)^*) \subset \mathcal{D}_{\overline{f}}. \) Fix \( z \in \mathcal{D}(\rho(f)^*), \) and let \( v = \rho(f)^*z. \) Now \( \rho(f_n) = \rho(f)\rho(f_n), \) and since \( \rho(f_n) \) is self-adjoint, \( \rho(f_n)\rho(f)^* = \rho(f)(\rho(f_n))^* = \rho(f_n)^* = \rho(\overline{f_n}), \) since \( f_n \) is bounded.

Thus \( \rho(f_n)v = \rho(\overline{f_n})z. \) Since \( |f_n| = 1, \)

\[
\int |f_n|^2 dE_{z,z} = \int |f_n| dE_{v,v} \cdot E_{v,v}(X) .
\]

Since this holds for every \( n, \) we see that \( z \in \mathcal{D}_{\overline{f}} = \mathcal{D}_{\overline{f}}. \) We leave the verification that \( \rho(f)^* = \rho(|f|^2) = \rho(f)^*\rho(f) \) to the reader.

### 24. The Spectral Theorem for Self-adjoint Operators

Before giving the Spectral Theorem, we present two results we will need. The first includes what is usually known as the Spectral Mapping Theorem (part c.) and will be used in the proof of the Spectral Theorem.

**Theorem** Let \( E \) be a resolution of the identity on \((X, \Omega, H), \) and let \( f \) be a measurable function on \( X. \) For \( \lambda \in \mathbb{C}, \) let \( \omega_{f\lambda} = f^{-1}(\lambda). \) Then

a. If \( \lambda \) is in the essential range of \( f \) and \( E(\omega_{f\lambda}) \cdot 0, \) then \( \rho(f) - \lambda I \) is not 1-1. That is, \( \lambda \) is an eigenvalue of \( \rho(f). \)
b. If $\lambda$ is in the essential range of $f$ and $E(\omega_{\lambda}) = 0$, then $\rho(f) - \lambda I$ is 1-1 and maps $\mathcal{D}_f$ onto a proper dense subspace of $H$. For any $\varepsilon > 0$, there exists a unit vector $x \in \mathcal{D}_f$ with $\|\rho(f)x - \lambda x\| < \varepsilon$.

c. $\sigma(\rho(f))$ = essential range of $f$.

**proof:** Parts a and b imply that $\text{EssentialRange}[f] \subset \sigma(\rho(f))$. If $\lambda_0 \notin \text{EssentialRange}[f]$, then $g = (f - \lambda_0) \in B(X)$ so $\rho(f)\rho(g) = \rho(1) = I$. Thus $\mathcal{R}(\rho(f)) = H$, from which we have $\rho(f)^{-1} \in B(H)$, so $\lambda_0 \notin \sigma(\rho(f))$.

As for parts a and b, let us assume $\lambda = 0$. If $E(\omega_0) \cdot 0$, choose $x_0 \cdot 0, x_0 \in \mathcal{R}(E(\omega_0))$. Since $f\chi_{\omega_0} = 0$ we have $\rho(f)\rho(\chi_{\omega_0}) = 0$. This implies $\rho(f)x_0 = (\rho(f)E(\omega_0))x_0 = 0$, so $\rho(f)$ is not 1-1.

Supposing $E(\omega_0) = 0$, we show first that $\rho(f)$ and $\rho(f)^* = \rho(\overline{f})$ are 1-1. The arguments are the same. If $\rho(f)x = 0$, then $\int |f|^2 dE_{x,x} = \langle \rho(f)x, x \rangle = 0$. But $|f| > 0$ almost everywhere $(dE_{x,x})$ so the vanishing of this integral implies that $E_{x,x}(X) = 0$. However $E_{x,x} = \|x\|^2$, so $x = 0$, and $\rho(f)$ is 1-1.

If now $y \in \mathcal{R}(\rho(f))^\perp$, then $x \rightarrow \langle \rho(f)x, y \rangle = 0$ is certainly a continuous linear functional. Thus $y \in \mathcal{D}(\rho(f)^*)$ and $\rho(f)^*y = 0$. Since $\rho(f)^*$ is 1-1, $y$ must be zero. Thus $\mathcal{R}(\rho(f))$ is dense. To show it is proper requires that we show $\rho(f)^{-1}$ is unbounded. Since $0 \in \text{EssentialRange}[f]$, with $\omega_\varepsilon = \{t: |f(t)| < \varepsilon\}$, we have $E(\omega_\varepsilon) \cdot 0$ for every positive $\varepsilon$. Choosing $x$ to be a unit vector in $\mathcal{R}(E(\omega_\varepsilon))$, we have $\|\rho(f)x\| = \|\rho(f)\chi_{\omega_\varepsilon})x\| \cdot \|\rho(f)\chi_{\omega_\varepsilon})\| < \varepsilon$. This shows $\rho(f)^{-1}$ is unbounded and completes the proof of b.

We will also need to transport the resolution of the identity on $T$ for the Cayley transform of $T$ to a resolution of the identity on $\sigma(T) \subset \mathbb{R}$. If $E$ is a resolution of the identity on $(X, \Omega, H)$, then the formula $E'(\omega') = E(\phi^{-1}(\omega'))$ defines a resolution of the identity on $(X', \Omega', H')$, for which we have $\int fdE'_{x,y} = \int (f\phi)dE_{x,y}$. We assume, of course that $\phi: X \rightarrow X'$ is measurable and 1-1.
**Spectral Theorem** Let $T$ be a self-adjoint operator on $H$. There exists a unique resolution of the identity on the Borel sets of $\sigma(T)$ such that

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

**proof:** Let $X' = T \setminus \{1\} = \{ t \in \mathbb{C} : t \cdot 1 \text{ and } |t| = 1 \}$, let $U$ be the Cayley transform of $T$, and let $E'$ be the resolution of the identity for $U$. Since $I - U$ is 1-1, $E'((1)) = 0$, and therefore $\langle UX, y \rangle = \int \lambda dE'_{X,Y}$. Now let $f(\lambda) = i \frac{1 + \lambda}{1 - \lambda}$. Then $\rho(f)x, y = \int f \, dE'_{x,y}$, and since $f$ is real valued on $T$, $\rho(f)$ is self-adjoint. Now $f(\lambda)(1 - \lambda) = i (1 + \lambda)$, and $\mathcal{R}(I - U) \subset D(\rho(f))$. But $T (I - U) = i (I + U)$ and $D(T) = \mathcal{R}(I - U)$, so $\rho(f)$ is a symmetric extension of $T$. But $T$ is self-adjoint, and hence maximally symmetric, so $T = \rho(f)$, and therefore $T = \int f \, dE'$. Now $\sigma(T)$ is the essential range of $f(\lambda)$, which is a subset of the real line. Since $f$ maps $X'$ 1-1 onto $\mathbb{R}$, we may define $E(f(\omega)) = E'/(\omega)$ to obtain

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

The uniqueness of this spectral representation is a consequence of the uniqueness of the integral representation for $U$.

**Exercises**

1. Let $T$ be a self adjoint operator on $H$. Show that $\langle Tx, x \rangle \geq 0 \forall x \in D(T) \iff \sigma(T) \subset \mathbb{R}$.

2. Show that if $T$ is a self-adjoint positive operator on $H$ then there is a unique self-adjoint operator $S$ on $H$ such that $S^2 = T$. 