The Spectral Theorem for Unbounded Operators.

Math 212a

November 29, 2001

Many important operators in Hilbert space that arise in physics and mathematics are "unbounded". For example the operator $D = \frac{1}{i}\frac{d}{dx}$ on $L_2(\mathbf{R})$. This operator is not defined on all of L_2, and where it is defined it is not bounded as an operator. One of the great achievements of Wintner in the late 1920's, followed by Stone and von Neumann was to prove a version of the spectral theorem for unbounded self-adjoint operators.

There are two (or more) approaches we could take to the proof of this theorem. Both involve the resolvent

$$R_z = R(z,T) := (zI - T)^{-1}. \tag{1}$$

After spending some time explaining what an unbounded operator is and giving the very subtle definition of what an unbounded self-adjoint operator is, we will prove that the resolvent of a self-adjoint operator exists and is a bounded normal operator for all non-real z.

We then apply the spectral theorem for bounded normal operators to derive the spectral theorem for unbounded self-adjoint operators. This is the fastest approach, but depends on the whole machinery of the Gelfand representation theorem.

As an alternative, I will also present a proof the spectral theorem for unbounded self-adjoint operators directly using (a mild modification of) Stone's formula. So I will present both methods. In the second method I will follow the treatment by Lorch in his little book *Spectral Theory* Oxford University Press (1962).

Contents

1 Operators and their domains.

Let B and C be Banach spaces. We make $B \oplus C$ into a Banach space via

$$\|\{x, y\}\| = \|x\| + \|y\|.$$

Here we are using $\{x, y\}$ to denote the ordered pair of elements $x \in B$ and $y \in C$ so as to avoid any conflict with our notation for scalar product in a Hilbert space. So $\{x, y\}$ is just another way of writing $x \oplus y$.

A linear subspace

$$\Gamma \subset B \oplus C$$

will be called a **graph** (more precisely a graph of a linear transformation) if

$$\{0, y\} \in \Gamma \quad \Rightarrow \quad y = 0.$$

Another way of saying the same thing is

$$\{x, y_1\} \in \Gamma \ \text{ and } \ \{x, y_2\} \in \Gamma \quad \Rightarrow \quad y_1 = y_2.$$

In other words, if $\{x, y\} \in \Gamma$ then y is determined by x. So let

$D(\Gamma)$ denote the set of all $x \in B$ such that there is a $y \in C$ with $\{x, y\} \in \Gamma$.

Then $D(\Gamma)$ is a linear subspace of B, but, and this is very important, $D(\Gamma)$ is *not* necessarily a closed subspace. We have a linear map

$$T(\Gamma): \ D(\Gamma) \to C, \quad Tx = y \ \text{ where } \ \{x, y\} \in \Gamma.$$

Equally well, we could start with the linear transformation: Suppose we are given a (not necessarily closed) subspace $D(T) \subset B$ and a linear transformation

$$T : D(T) \to C.$$

We can then consider its graph $\Gamma(T) \subset B \oplus C$ which consists of all

$$\{x, Tx\}.$$

Thus the notion of a graph, and the notion of a linear transformation defined only on a subspace of B are logically equivalent. When we start with T (as usually will be the case) we will write $D(T)$ for the domain of T and $\Gamma(T)$ for the corresponding graph. There is a certain amount of abuse of language here, in that when we write T, we mean to include $D(T)$ and hence $\Gamma(T)$ as part of the definition.

A linear transformation is said to be **closed** if its graph is a closed subspace of $B \oplus C$. Let us disentangle what this says for the operator T. It says that if $f_n \in D(T)$ then

$$f_n \to f \text{ and } Tf_n \to g \quad \Rightarrow \quad f \in D(T) \text{ and } Tf = g.$$

This is a much weaker requirement than continuity. Continuity of T would say that $f_n \to f$ alone would imply that Tf_n converges to Tf. Closedness says that if we know that *both* f_n converges *and* that $g_n = Tf_n$ converges then we can conclude that $f = \lim f_n$ lies in $D(T)$ and that $Tf = g$.

An important theorem, known as the *closed graph theorem* says that if T is closed and $D(T)$ is all of B then T is bounded. As we will not need to use this theorem in this lecture, we will not present its proof here. For the proof, see the appendix to these notes.

2 The adjoint.

Suppose that we have a linear operator $T : D(T) \to C$ and let us make the hypothesis that

$$D(T) \textit{ is dense in } B.$$

Any element of B^* is then completely determined by its restriction to $D(T)$. Now consider

$$\Gamma(T)^* \subset C^* \oplus B^*$$

defined by

$$\{\ell, m\} \in \Gamma(T)^* \quad \Leftrightarrow \quad \langle \ell, Tx \rangle = \langle m, x \rangle \quad \forall\, x \in D(T). \tag{2}$$

Since m is determined by its restriction to $D(T)$, we see that $\Gamma^* = \Gamma(T^*)$ is indeed a graph. (It is easy to check that it is a linear subspace of $C^* \oplus B^*$.) In other words we have defined a linear transformation

$$T^* := T(\Gamma(T)^*)$$

whose domain consists of all $\ell \in C^*$ such that there exists an $m \in B^*$ for which $\langle \ell, Tx \rangle = \langle m, x \rangle \quad \forall\, x \in D(T)$.

If $\ell_n \to \ell$ and $m_n \to m$ then the definition of convergence in these spaces implies that for any $x \in D(T)$ we have

$$\langle \ell, Tx \rangle = \lim \langle \ell_n, Tx \rangle = \lim \langle m_n, x \rangle = \langle m, x \rangle.$$

If we let x range over all of $D(T)$ we conclude that Γ^* is a closed subspace of $C^* \oplus B^*$. In other words we have proved

Theorem 1 *If $T : D(T) \to C$ is a linear transformation whose domain $D(T)$ is dense in B, it has a well defined adjoint T^* whose graph is given by (2). Furthermore T^* is a closed operator.*

3 Self-adjoint operators.

Now let us restrict to the case where $B = C = H$ is a Hilbert space, so we may identify $B^* = C^* = H^*$ with H via the Riesz representation theorem. If $T : D(T) \to H$ is an operator with $D(T)$ dense in H we may identify the graph of T^* as consisting of all $\{g, h\} \in H \oplus H$ such that

$$(Tx, g) = (x, h) \quad \forall x \in D(T)$$

and then write

$$(Tx, g) = (x, T^*g) \quad \forall\, x \in D(T), \quad g \in D(T^*).$$

We now come to the central definition: An operator A defined on a domain $D(A) \subset H$ is called **self-adjoint** if

- $D(A)$ is dense in H,
- $D(A) = D(A^*)$, and
- $Ax = A^*x \;\forall x \in D(A)$.

The conditions about the domain $D(A)$ are rather subtle, and we shall go into some of their subtleties in a later lecture. For the moment we record one immediate consequence of the theorem of the preceding section:

Proposition 1 *Any self adjoint operator is closed.*

If we combine this proposition with the closed graph theorem which asserts that a closed operator defined on the whole space must be bounded, we derive a famous theorem of Hellinger and Toeplitz which asserts that any self adjoint operator defined on the whole Hilbert space must be bounded. This shows that for self-adjoint operators, being globally defined and being bounded amount to the same thing. At the time of its appearance in the second decade of the twentieth century, this theorem of Hellinger and Toeplitz was considered an astounding result. It was only after the work of Banach, in particular the proof of the closed graph theorem, that this result could be put in proper perspective.

4 The resolvent.

The following theorem will be central for us.

Theorem 2 *Let A be a self-adjoint operator on a Hilbert space H with domain $D = D(A)$. Let*

$$c = \lambda + i\mu, \quad \mu \neq 0$$

be a complex number with non-zero imaginary part. Then

$$(cI - A) : D(A) \to H$$

is bijective. Furthermore the inverse transformation

$$(cI - A)^{-1} : H \to D(A)$$

is bounded and in fact

$$\|(cI - A)^{-1}\| \leq \frac{1}{|\mu|}. \tag{3}$$

Remark. In the case of a *bounded* self adjoint operator this is an immediate consequence of the spectral theorem, more precisely of the fact that Gelfand transform is an isometric isomorphism of the closed algebra generated by A with the algebra $C(\operatorname{Spec} A)$. Indeed, the function $x \mapsto 1/(c - x)$ is bounded on the whole real axis with supremum $1/|\mu|$. Since $\operatorname{Spec}(A) \subset \mathbf{R}$ we conclude that $(cI - A)^{-1}$ exists and its norm satisfies (3).

We also gave a direct proof in class of Theorem 2 for *bounded* self-adjoint operators.

We will now give a proof of this theorem valid for general self-adjoint operators, and will use this theorem for the proof of the spectral theorem in the general case. At certain critical junctures, you will see where the subtle condition $D(A) = D(A^*)$ comes in.

Proof. Let $g \in D(A)$ and set

$$f = (cI - A)g = [\lambda I - A]g + i\mu g.$$

Then $\|f\|^2 = (f, f) =$

$$\|[cI - A]g\|^2 + \mu^2\|g\|^2 + ([\lambda I - A]g, i\mu g) + (i\mu g, [\lambda I - A]g).$$

I claim that these last two terms cancel. Indeed, since $g \in D(A)$ and A is self adjoint we have $g \in D(A^*$ and

$$(\mu g, [\lambda I - A]g) = (\mu[\lambda I - A]g, g) = ([\lambda I - A]g, \mu g)$$

since μ is real. Hence

$$([\lambda I - A]g, i\mu g) = -i(\mu g, [\lambda I - A]g) = -(i\mu g, [\lambda I - Ag).$$

We have thus proved that

$$\|f\|^2 = \|(\lambda I - A)g\|^2 + \mu^2\|g\|^2. \tag{4}$$

In particular

$$\|f\|^2 \geq \mu^2\|g\|^2$$

for all $g \in D(A)$. Since $|\mu| > 0$, we see that $f = 0 \Rightarrow g = 0$ so $(cI - A)$ is injective on $D(A)$, and furthermore that that $(cI - A)^{-1}$ (which is defined on the image of $(cI - A)$)satisfies (3). We must show that this image is all of H.

First we show that the image is dense. For this it is enough to show that there is no $h \neq 0 \in H$ which is orthogonal to the image of $(cI - A)$. So suppose that

$$([cI - A]g, h) = 0 \quad \forall g \in D(A).$$

Then

$$(g, \overline{c}h) = (cg, h) = (Ag, h) \quad \forall g \in D(A)$$

which says that $h \in D(A^*)$ and $A^*h = \overline{c}h$. But A is self adjoint so $h \in D(A)$ and $Ah = \overline{c}h$. Thus

$$\overline{c}(h, h) = (\overline{c}h, h) = (Ah, h) = (h, Ah) = (h, \overline{c}h) = c(h, h).$$

Since $c \neq \overline{c}$ this is impossible unless $h = 0$. We have now established that the image of $cI - A$ is dense in H. Notice that in this argument we used crucially the assumption that $D(A^*) = D(A)$.

We now prove that the image of $(cI - A)$ is all of H. So let $f \in H$. We know that we can find

$$f_n = (cI - A)g_n, \quad g_n \in D(A) \quad \text{with } f_n \to f.$$

The sequence f_n is convergent, hence Cauchy, and from (3) applied to elements of $D(A)$ we know that

$$\|g_m - g_n\| \leq |\mu|^{-1}\|f_n - f_m\|.$$

Hence the sequence $\{g_n\}$ is Cauchy, so $g_n \to g$ for some $g \in H$. But we know by Proposition 1 that A is a closed operator. Hence $g \in D(A)$ and $(cI - A)g = f$. QED

We repeat that the operator

$$R_z = R_z(A) = (zI - A)^{-1}$$

is called the **resolvent** of A when it exists as a bounded operator. The set of $z \in \mathbf{C}$ for which the resolvent exists is called the **resolvent set** and the complement of the resolvent set is called the **spectrum** of the operator. For bounded self-adjoint operators this terminology coincides with the terminology we introduced earlier.

The preceding theorem asserts that the spectrum of a self-adjoint operator is a subset of the real numbers.

Let z and w both belong to the resolvent set. We have

$$wI - A = (w - z)I + (zI - A).$$

Multiplying this equation on the left by R_w gives

$$I = (w - z)R_w + R_w(zI - A),$$

and multiplying this on the right by R_z gives

$$R_z - R_w = (w - z)R_w R_z.$$

From this it follows (by interchanging z and w) that $R_z R_w = R_w R_z$, in other words all resolvents R_z commute with one another and we can also write the preceding equation as

$$R_z - R_w = (w - z)R_z R_w. \tag{5}$$

This equation, which is known as the **resolvent equation** dates back to the theory of integral equations in the nineteenth century.

It follows from the resolvent equation that $z \mapsto R_z$ (for fixed A) is a continuous function of z. Once we know that the resolvent is a continuous function of z, we may divide the resolvent equation by $(z - w)$ if $z \neq w$ and, if w is interior to the resolvent set, conclude that

$$\lim_{z \to w} \frac{R_z - R_w}{z - w} = -R_w^2.$$

This says that the "derivative in the complex sense" of the resolvent exists and is given by $-R_z^2$. In other words, the resolvent is a "holomorphic operator valued" function of z.

To emphasize this holomorphic character of the resolvent, we have

Proposition 2 *Let z belong to the resolvent set. The the open disk of radius $\|R_z\|^{-1}$ about z belongs to the resolvent set and on this disk we have*

$$R_w = R_z(I + (z - w)R_z + (z - w)^2 R_z^2 + \cdots). \tag{6}$$

Proof. The series on the right converges in the uniform topology since $|z - w| < \|R_z\|^{-1}$. Multiplying this series by $(zI - A) - (z - w)I$ gives I. But $zI - A - (z - w)I = wI - A$. So the right hand side is indeed R_w. QED

This suggests that we can develop a "Cauchy theory" of integration of functions such as the resolvent, and we shall do so, eventually leading to a proof of the spectral theorem for unbounded self-adjoint operators.

However we first give a proof (following the treatment in Reed-Simon) in which we derive the spectral theorem for unbounded operators from the Gelfand representation theorem applied to the closed algebra generated by the *bounded* normal operators $(\pm iI - A)^{-1}$.

5 The multiplication operator form of the spectral theorem.

We first state this theorem for closed commutative self-adjoint algebras of (bounded) operators. Recall that "self-adjoint" in this context means that if $T \in B$ then $T^* \in B$.

Theorem 3 *Let B be a commutative closed self-adjoint subalgebra of the algebra of all bounded operators on a separable Hilbert space H. Then there exists a measure space $(M, \mathcal{F}, \mu)$ with $\mu(M) < \infty$, a unitary isomorphism*

$$W : H \to L_2(M, \mu),$$

and a map

$$B \to \quad \text{bounded measurable functions on } M, \qquad T \mapsto \tilde{T}$$

such that

$$[(WTW^{-1})f](m) = \tilde{T}(m)f(m).$$

In fact, M can be taken to be a finite or countable disjoint union of $\mathcal{M} = \mathrm{Mspec}(B)$

$$M = \bigcup_1^N \mathcal{M}_i, \qquad \mathcal{M}_i = \mathcal{M}$$

$N \in \mathbf{Z}_+ \cup \infty$ and

$$\tilde{T}(m) = \hat{T}(m) \quad \text{if } m \in \mathcal{M}_i = \mathcal{M}.$$

In short, the theorem says that any such B is isomorphic to an algebra of multiplication operators on an L_2 space. We prove the theorem in two stages.

5.1 Cyclic vectors.

An element $x \in H$ is called a **cyclic vector** for B if $Bx = H$. In more mundane terms this says that the space of linear combinations of the vectors Tx, $T \in B$ are dense in H.

For example, if B consists of all multiples of the identity operator, then Bx consists of all multiples of x, so B can not have a cyclic vector unless H is one dimensional. More generally, if H is finite dimensional and B is the algebra generated by a self-adjoint operator, then B can not have a cyclic vector if A has a repeated eigenvalue.

Proposition 3 *Suppose that x is a cyclic vector for B. Then it is a cyclic vector for the projection valued measure P on $\mathcal{M}$ associated to B in the sense that the linear combinations of the vectors $P(U)x$ are dense in H as U ranges over the Borel sets on $\mathcal{M}$.*

Proof. Suppose not. Then there exists a non-zero $y \in H$ such that

$$(P(U)x, y) = 0$$

for all Borel subset U of $\mathcal{M}$. Then, in the terminology of the notes on Banach algebras,

$$(Tx, y) = \int_{\mathcal{M}} \hat{T} d(P(U)x, y) = \int_{\mathcal{M}} \hat{T} d\mu_{x,y} = 0,$$

which contradicts the assumption that the linear combinations of the Tx are dense in H. QED

Let us continue with the assumption that x is a cyclic vector for B. Let

$$\mu = \mu_{x,x}$$

so

$$\mu(U) = (P(U)x, x),$$

This is a finite measure on $\mathcal{M}$, in fact

$$\mu(\mathcal{M}) = \|x\|^2, \tag{7}$$

since the function $\mathbf{1} = \mathbf{1}_{\mathcal{M}}$ corresponds to the identity operator in the Gelfand representation. We will construct a unitary isomorphism of

$$W: \quad H \to L_2(\mathcal{M}, \mu)$$

starting with the assignment

$$Wx = \mathbf{1} = \mathbf{1}_{\mathcal{M}}.$$

We would like this to be a B morphism, even a morphism for the action of multiplication by bounded Borel functions. This forces the definition

$$WP(U)x = \mathbf{1}_U \mathbf{1} = \mathbf{1}_U.$$

This then forces

$$W[c_1 P(U_1)x + \cdots c_n P(U_n)x] = s$$

for any simple function

$$s = c_1 \mathbf{1}_{U_1} + \cdots c_n \mathbf{1}_{U_n}.$$

A direct check shows that this is well defined for simple functions. We can write this map as

$$W\left[O(s)x\right] = s,$$

and another direct check shows that

$$\|W\left[O(s)x\right]\| = \|s\|_2$$

where the norm on the right is the L_2 norm relative to the measure μ. Since the simple functions are dense in $L_2(\mathcal{M},\mu)$ and the vectors $O(s)x$ are dense in H this extends W to a unitary isomorphism of H onto $L_2(M,\mu)$. Furthermore,

$$W^{-1}(f) = O(f)x$$

for any $f \in L_2(\mathcal{M},\mu)$. For simple functions, and therefore for all $f \in L_2(\mathcal{M},\mu)$ we have

$$W^{-1}(\hat{T}f) = O(\hat{T}f)x = TO(f)x = TW^{-1}(f)$$

or

$$(WTW^{-1})f = \hat{T}f$$

which is the assertion of the theorem. In other words we have proved the theorem under the assumption of the existence of a cyclic vector.

5.2 The general case.

Start with any non-zero vector x_1 and consider $H_1 = Bx_1 =$ the closure of linear combinations of Tx_1, $T \in B$. The space H_1 is a closed subspace of H which is invariant under B, i.e. $TH_1 \subset H_1 \;\; \forall T \in B$. Therefore the space $H_1^{\perp}$ is also invariant under B since if $(x_1, y) = 0$ then

$$(x_1, Ty) = (T^* x_1, y) = 0 \;\text{ since }\; T^* \in B.$$

Now if $H_1 = H$ we are done, since x_1 is a cyclic vector for B acting on H_1. If not choose a non-zero $x_2 \in H_2$ and repeat the process. We can choose a collection of non-zero vectors z_i whose linear combinations are dense in H - this is the separability assumption. So we may choose our x_i to be obtained from orthogonal projections applied to the z_i. In other words we have

$$H = H_1 \oplus H_2 \oplus H_3 \oplus \cdots$$

where this is either a finite or a countable Hilbert space (completed) direct sum. Now let us also take care to choose our x_n so that

$$\sum \|x_n\|^2 < \infty$$

which we can do, since $c_n x_n$ is just as good as x_n for any $c_n \neq 0$. We have a unitary isomorphism of H_n with $L_2(\mathcal{M}, \mu_n)$ where $\mu_n(U) = (P(U)x_n, x_n)$. In particular,

$$\mu_n(\mathcal{M}) = \|x_n\|^2.$$

So if we take M to be the disjoint union of copies $\mathcal{M}_n$ of $\mathcal{M}$ each with measure μ_n then the total measure of M is finite and

$$L_2(M) = \bigoplus L_2(\mathcal{M}_n, \mu_n)$$

where this is either a finite direct sum or a (Hilbert space completion of) a countable direct sum. Thus the theorem for the cyclic case implies the theorem for the general case. QED

5.3 The spectral theorem for unbounded self-adjoint operators, multiplication operator form.

We now let A be a (possibly unbounded) self-adjoint linear transformation. Let us apply the previous theorem to the algebra generated by the bounded operators $(\pm iI - A)^{-1}$ which are the adjoints of one another. Observe that there is no non-zero vector $y \in H$ such that

$$(A + iI)^{-1} y = 0.$$

Indeed if such a $y \in H$ existed, we would have

$$0 = (x, (A + iI)^{-1} y) = ((A - iI)^{-1} x, y) = -((iI - A)^{-1} x, y) \quad \forall\, x \in H$$

and we know that the image of $(iI - A)^{-1}$ is $D(A)$ which is dense in H.

Now consider the function $\left((A + iI)^{-1}\right)\tilde{}$ on M given by Theorem 3. It can not vanish on any set of positive measure, since any function $f \neq 0$ supported on such a set would be in the kernel of the operator consisting of multiplication by $\left((A + iI)^{-1}\right)\tilde{}$, and hence $y = W^{-1} f$ would satisfy $(A + iI)^{-1} y = 0$.

Thus the function

$$\tilde{A} := \left[\left((A + iI)^{-1}\right)\tilde{}\,\right]^{-1} - i$$

is finite almost everywhere on M relative to the measure μ although it might (and generally will) be unbounded. Our plan is to show that under the unitary isomorphism W the operator A goes over into multiplication by $\tilde{A}$.

First we show

Proposition 4 *$x \in D(A)$ if and only if $\tilde{A} W x \in L_2(M, \mu)$.*

Proof. Suppose $x \in D(A)$. Then $x = (A + iI)^{-1} y$ for some $y \in H$ and so

$$Wx = \left((A + iI)^{-1}\right)\tilde{}\, f, \qquad f = Wy.$$

But

$$\tilde{A} \left((A + iI)^{-1}\right)\tilde{} = 1 - ih$$

where

$$h = \left((A+iI)^{-1}\right)\tilde{}$$

is a bounded function. Thus $\tilde{A}Wx \in L_2(M,\mu)$.

Conversely, if $\tilde{A}Wx \in L_2(M,\mu)$, then $(\tilde{A}+iI)Wx \in L_2(M,\mu)$, which means that there is a $y \in H$ such that $Wy = (\tilde{A}+iI)Wx$. Therefore

$$\left((A+iI)^{-1}\right)\tilde{}\,Wy = \left((A+iI)^{-1}\right)\tilde{}\,(\tilde{A}+iI)Wx = Wx$$

and hence

$$x = (A+iI)^{-1}y \in D(A).$$

QED

Proposition 5 *If $h \in W(D(A))$ then $\tilde{A}h = WAW^{-1}h$.*

Proof. Let $x = W^{-1}h$ which we know belongs to $D(A)$ so we may write $x = (A+iI)^{-1}y$ for some $y \in H$, and hence

$$Ax = y - ix \quad \text{and } Wy = \left[\left((A+iI)^{-1}\right)\tilde{}\,\right]^{-1} h.$$

So

$$\begin{aligned} WAx &= Wy - iWx \\ &= \left[\left((A+iI)^{-1}\right)\tilde{}\,\right]^{-1} h - ih \\ &= \tilde{A}h \quad \text{QED} \end{aligned}$$

The function $\tilde{A}$ must be real valued almost everywhere since if its imaginary part were positive (or negative) on a set U of positive measure, then $(\tilde{A}\mathbf{1}_U, \mathbf{1}_U)_2$ would have non-zero imaginary part contradicting the fact that multiplication by $\tilde{A}$ is a self adjoint operator, being unitarily equivalent to the self adjoint operator A.

Putting all this together we get

Theorem 4 *Let A be a self adjoint operator on a separable Hilbert space H. Then there exists a finite measure space (M,μ), a unitary isomorphism $W : H \to L_2(M,\mu)$ and a real valued measurable function $\tilde{A}$ which is finite almost everywhere such that $x \in D(A)$ if and only if $\tilde{A}Wx \in L_2(M,\mu)$ and if $h \in W(D(A))$ then $\tilde{A}h = WAW^{-1}h$.*

5.4 The functional calculus.

Let f be any bounded Borel function defined on $\mathbf{R}$. Then

$$f \circ \tilde{A}$$

is a bounded function defined on M. Multiplication by this function is a bounded operator on $L_2(M,\mu)$ and hence corresponds to a bounded self-adjoint operator on H. With a slight abuse of language we might denote this operator by $O(f \circ \tilde{A})$. However we will use the more suggestive notation

$$f(A).$$

The map

$$f \mapsto f(A)$$

- is an algebraic homomorphism,
- $\overline{f}(A) = f(A)^*$,
- $\|f(A)\| \leq \|f\|_\infty$ where the norm on the left is the uniform operator norm and the norm on the right is the sup norm on $\mathbf{R}$
- if $Ax = \lambda x$ then $f(A)x = f(\lambda)x$,
- if $f \geq 0$ then $f(A) \geq 0$ in the operator sense,
- if $f_n \to f$ pointwise and if $\|f_n\|_\infty$ is bounded, then $f_n(A) \to f(A)$ strongly, and
- if f_n is a sequence of Borel functions on the line such that $|f_n(\lambda)| \leq |\lambda|$ for all n and for all $\lambda \in \mathbf{R}$, and if $f_n(\lambda) \to \lambda$ for each fixed $\lambda \in \mathbf{R}$ then for each $x \in D(A)$
$$f_n(A)x \to Ax.$$

All of the above statements are obvious except perhaps for the last two which follow from the dominated convergence theorem. It is also clear from the preceding discussion that the map $f \mapsto f(A)$ is uniquely determined by the above properties.

Multiplication by the function $e^{it\tilde{A}}$ is a unitary operator on $L_2(M, \mu)$ and

$$e^{is\tilde{A}} e^{it\tilde{A}} = e^{i(s+t)\tilde{A}}.$$

Hence from the above we conclude

Theorem 5 **[Half of Stone's theorem.]** *For an self adjoint operator A the operator e^{itA} given by the functional calculus as above is a unitary operator and*

$$t \mapsto e^{itA}$$

is a one parameter group of unitary transformations.

The full Stone's theorem asserts that any unitary one parameter groups is of this form. We will discuss this later.

5.5 Resolutions of the identity

For each measurable subset X of the real line we can consider its indicator function $\mathbf{1}_X$ and hence $\mathbf{1}_X(A)$ which we shall denote by $P(X)$. In other words

$$P(X) := \mathbf{1}_X(A).$$

It follows from the above that

$$\begin{aligned}
P(X)^* &= P(X) \\
P(X)P(Y) &= P(X \cap Y) \\
P(X \cup Y) &= P(X) + P(Y) \quad \text{if} \quad X \cap Y = 0 \\
P(X) &= s - \lim \sum_1^N P(X_i) \quad \text{if } X_i \cap X_j = \emptyset \text{ if } i \neq j \quad \text{and } X = \bigcup X_i \\
P(\emptyset) &= 0 \\
P(\mathbf{R}) &= I.
\end{aligned}$$

For each $x, y \in H$ we have the complex valued measure

$$P_{x,y}(X) = (P(X)x, y)$$

and for any bounded Borel function f we have

$$(f(A)x, y) = \int_{\mathbf{R}} f(\lambda) dP_{x,y}.$$

If g is an unbounded (complex valued) Borel function on $\mathbf{R}$ we define $D(g(A))$ to consist of those $x \in H$ for which

$$\int_{\mathbf{R}} |g(\lambda)|^2 dP_{x,x} < \infty.$$

The set of such x is dense in H and we define $g(A)$ on $D(A)$ by

$$(g(A)x, y) = \int_{\mathbf{R}} g(\lambda) dP_{x,y}$$

for $x, y \in D(g(A))$ (and the Riesz representation theorem). This is written symbolically as

$$g(A) = \int_{\mathbf{R}} g(\lambda) dP.$$

In the special case $g(\lambda) = \lambda$ we write

$$A = \int_{\mathbf{R}} \lambda dP.$$

this is the spectral theorem for self adjoint operators.

In the older literature one often sees the notation

$$E_\lambda := P(-\infty, \lambda).$$

A translation of the properties of P into properties of E is

$$\begin{aligned} E_\lambda^2 &= E_\lambda & (8)\\ E_\lambda^* &= E_\lambda & (9)\\ \lambda < \mu &\Rightarrow E_\lambda E_\mu = E_\lambda & (10)\\ \lambda_{n\rightarrow} - \infty &\Rightarrow E_{\lambda_n} \rightarrow 0 \text{ strongly} & (11)\\ \lambda_n \rightarrow +\infty &\Rightarrow E_{\lambda_n} \rightarrow I \text{ strongly} & (12)\\ \lambda_n \nearrow \lambda &\Rightarrow E_n \rightarrow E_\lambda \text{ strongly.} & (13) \end{aligned}$$

One then writes the spectral theorem as

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda. \tag{14}$$

We shall now give an alternative proof of this formula which does not depend on either the Gelfand representation theorem of any of the limit theorems of Lebesgue integration. Instead, it depends on the Riesz-Dunford extension of the Cauchy theory of integration of holomorphic functions along curves to operator valued holomorphic functions.

6 The Riesz-Dunford calculus.

Suppose that we have a continuous map $z \mapsto S_z$ defined on some open set of complex numbers, where S_z is a bounded operator on some fixed Banach space and by continuity, we mean continuity relative to the uniform metric on operators. If C is a continuous piecewise differentiable (or more generally any rectifiable) curve lying in this open set, and if $t \mapsto z(t)$ is a piecewise smooth (or rectifiable) parametrization of this curve, then the map $t \mapsto S_{z(t)}$ is continuous.

For any partition $0 = t_0 \leq t_1 \leq \cdots \leq t_n = 1$ of the unit interval we can form the Cauchy approximating sum

$$\sum_{i=1}^{n} S_{z(t_i)}(z(t_i) - z(t_{i-1})),$$

and the usual proof of the existence of the Riemann integral shows that this tends to a limit as the mesh becomes more and more refined and the mesh distance tends to zero. The limit is denoted by

$$\int_C S_z dz$$

and this notation is justified because the change of variables formula for an ordinary integral shows that this value does not depend on the parametrization, but only on the orientation of the curve C.

We are going to apply this to $S_z = R_z$, the resolvent of an operator, and the main equations we shall use are the resolvent equation (5) and the power series for the resolvent (6). We repeat these equations here:

$$R_z - R_w = (w - z)R_z R_w$$

and

$$R_w = R_z(I + (z - w)R_z + (z - w)^2 R_z^2 + \cdots).$$

We proved that the resolvent of a self-adjoint operator exists for all non-real values of z.

But a lot of the theory goes over for the resolvent

$$R_z = R(z, T) = (zI - T)^{-1}$$

where T is an arbitrary operator on a Banach space, so long as we restrict ourselves to the resolvent set, i.e. the set where the resolvent exists as a bounded operator. So, following Lorch *Spectral Theory* we first develop some facts about integrating the resolvent in the more general Banach space setting (where our principal application will be to the case where T is a bounded operator).

For example, suppose that C is a simple closed curve contained in the disk of convergence about z of (6) i.e. of the above power series for R_w. Then we can integrate the series term by term. But

$$\int_C (z - w)^n dw = 0$$

for all $n \neq -1$ so

$$\int_C R_w dw = 0.$$

By the usual method of breaking any any deformation up into a succession of small deformations and then breaking any small deformation up into a sequence of small "rectangles" we conclude

Theorem 6 *If two curves C_0 and C_1 lie in the resolvent set and are homotopic by a family C_t of curves lying entirely in the resolvent set then*

$$\int_{C_0} R_z dz = \int_{C_1} R_z dz.$$

Here are some immediate consequences of this elementary result.

Suppose that T is a bounded operator and $|z| > \|T\|$. Then

$$(zI - T)^{-1} = z^{-1}(I - z^{-1}T)^{-1} = z^{-1}(I + z^{-1}T + z^{-2}T^2 + \cdots)$$

exists because the series in parentheses converges in the uniform metric. In other words, all points in the complex plane outside the disk of radius $\|T\|$ lie in the resolvent set of T. From this it follows that the spectrum of any bounded operator can not be empty (if the Banach space is not $\{0\}$). (Recall the the spectrum is the complement of the resolvent set.) Indeed, if the resolvent set were the whole plane, then the circle of radius zero about the origin would be homotopic to a circle of radius $> \|T\|$ via a homotopy lying entirely in the resolvent set. Integrating R_z around the circle of radius zero gives 0. We can integrate around a large circle using the above power series. In performing this integration, all terms vanish except the first which give $2\pi i I$ by the usual Cauchy integral (or by direct computation). Thus $2\pi I = 0$ which is impossible in a non-zero vector space.

Here is another very important (and easy) consequence of the preceding theorem:

Theorem 7 *Let C be a simple closed rectifiable curve lying entirely in the resolvent set of T. Then*

$$P := \frac{1}{2\pi i}\int_C R_z dz \tag{15}$$

is a projection which commutes with T, i.e.

$$P^2 = P \quad \text{and} \quad PT = TP.$$

Proof. Choose a simple closed curve disjoint from C but sufficiently close to C so as to be homotopic to C via a homotopy lying in the resolvent set. Thus

$$P = \frac{1}{2\pi i}\int_{C'} R_w dw$$

and so

$$(2\pi i)^2 P^2 = \int_C R_z dz \int_{C'} R_w dw = \int_C \int_{C'} (R_w - R_z)(z-w)^{-1} dw dz$$

where we have used the resolvent equation (5). We write this last expression as a sum of two terms,

$$\int_{C'} R_w \int_C \frac{1}{z-w} dz dw - \int_C R_z \int_{C'} \frac{1}{z-w} dw dz.$$

Suppose that we choose C' to lie entirely inside C. Then the first expression above is just $(2\pi i)\int_{C'} R_w dw$ while the second expression vanishes, all by the elementary Cauchy integral of $1/(z-w)$. Thus we get

$$(2\pi i)^2 P^2 = (2\pi i)^2 P$$

or $P^2 = P$. This proves that P is a projection. It commutes with T because it is an integral whose integrand R_z commutes with T for all z. QED

The same argument proves

Theorem 8 *Let C and C' be simple closed curves each lying in the resolvent set, and let P and P' be the corresponding projections given by (15). Then $PP' = 0$ if the curves lie exterior to one another while $PP' = P'$ if C' is interior to C.*

Let us write

$$B' := PB, \quad B'' = (I-P)B$$

for the images of the projections P and $I-P$ where P is given by (15). Each of these spaces is invariant under T and hence under R_z because $PT = TP$ and hence $PR_z = R_zP$.

For any transformation C commuting with P let us write

$$C' := PC = CP = PCP \quad \text{and } C'' = (I-P)C = C(I-P) = (I-P)C(I-P)$$

so that C' and C'' are the restrictions of C to B' and B'' respectively.

For example, we may consider $R'_z = PR_z = R_zP$. For $x' \in B'$ we have $R'_z(zI - T')x' = R_zP(zI - TP)x' = R_z(zI-T)Px' = x'$. In other words R'_z is the resolvent of T' (on B') and similarly for R''_z. So if z is in the resolvent set for T it is in the resolvent set for T' and T''.

Conversely, suppose that z is in the resolvent set for both T' and T''. Then there exists an inverse A_1 for $zI' - T'$ on B' and an inverse A_2 for $zI'' - T''$ on B'' and so $A_1 \oplus A_2$ is the inverse of $zI - T$ on $B = B' \oplus B''$.

So a point belongs to the resolvent set of T if and only if it belongs to the resolvent set of T' and of T''. Since the spectrum is the complement of the resolvent set, we can say that a point belongs to the spectrum of T if and only if it belongs either to the spectrum of T' or of T'':

$$\text{Spec}(T) = \text{Spec}(T') \cup \text{Spec}(T'').$$

We now show that this decomposition is in fact the decomposition of Spec(T) into those points which lie inside C and outside C.

So we must show that if z lies exterior to C then if lies in the resolvent set of T'. This will certainly be true if we can find a transformation A on B which commutes with T and such that

$$A(zI - T) = P \tag{16}$$

for then A' will be the resolvent at z of T'. Now

$$(zI-T)R_w = (wI-T)R_w + (z-w)R_w = I + (z-w)R_w$$

so

$$(zI-T)\cdot\frac{1}{2\pi i}\int_C R_w \cdot \frac{1}{z-w}dw =$$

$$= \frac{1}{2\pi i}\int_C \frac{1}{z-w}dw \cdot I + \frac{1}{2\pi i}\int_C R_w dw = 0 + P = P.$$

We have thus proved

Theorem 9 *Let T be a bounded linear transformation on a Banach space and C a simple closed curve lying in its resolvent set. Let P be the projection given by (15) and*

$$B = B' \oplus B'', \quad T = T' \oplus T''$$

the corresponding decomposition of B *and of* T. *Then* $\mathrm{Spec}(T')$ *consists of those points of* $\mathrm{Spec}(T)$ *which lie inside* C *and* $\mathrm{Spec}(T'')$ *consists of those points of* $\mathrm{Spec}(T)$ *which lie exterior to* C.

We now begin to have a better understanding of Stone's formula: Suppose A is a self-adjoint operator. We know that its spectrum lies on the real axis. If we draw a rectangle whose upper and lower sides are parallel to the axis, and if its vertical sides do not intersect $\mathrm{Spec}(A)$, we would get a projection onto a subspace M of our Hilbert space which is invariant under A, and such that the spectrum of A when restricted to M lies in the interval cut out on the real axis by our rectangle. The problem is how to make sense of this procedure when the vertical edges of the rectangle might cut through the spectrum, in which case the integral (15) might not even be defined. This is resolved by the method of Lorch (the exposition is taken from his book) which we explain in the next section.

7 Lorch's proof of the spectral theorem.

7.1 Positive operators

Recall that if A is a bounded self-adjoint operator on a Hilbert space H then we write $A \geq 0$ if $(Ax, x) \geq 0$ for all $x \in H$ and (by a slight abuse of language) call such an operator positive. Clearly the sum of two positive operators is positive as is the multiple of a positive operator by a non-negative number. Also we write $A_1 \geq A_2$ for two self adjoint operators if $A_1 - A_2$ is positive.

Proposition 6 *If* A *is a bounded self-adjoint operator and* $A \geq I$ *then* A^{-1} *exists and*
$$\|A^{-1}\| \leq 1.$$

Proof. We have
$$\|Ax\| \, \|x\| \geq (Ax, x) \geq (x, x) = \|x\|^2$$
so
$$\|Ax\| \geq \|x\| \quad \forall \; x \in H.$$
So A is injective, and hence A^{-1} is defined on im A and is bounded by 1 there. We must show that this image is all of H.

If y is orthogonal to im A we have
$$(x, Ay) = (Ax, y) = 0 \;\; \forall \; x \in H$$
so $Ay = 0$ so $(y, y) \leq (Ay, y) = 0$ and hence $y = 0$. Thus im A is dense in H.

Suppose that $Ax_n \to z$. Then the x_n form a Cauchy sequence by the estimate above on $\|A^{-1}\|$ and so $x_n \to x$ and the continuity of A implies that $Ax = z$. QED

Suppose that $A \geq 0$. Then for any $\lambda > 0$ we have $A + \lambda I \geq \lambda I$, and by the proposition $(A + \lambda I)^{-1}$ exists, i.e. $-\lambda$ belongs to the resolvent set of A. So we have proved.

Proposition 7 *If* $A \geq 0$ *then* $\mathrm{Spec}(A) \subset [0, \infty)$.

Theorem 10 *If A is a self-adjoint transformation then*

$$\|A\| \leq 1 \quad \Leftrightarrow \quad -I \leq A \leq I. \tag{17}$$

Proof. Suppose $\|A\| \leq 1$. Then using Cauchy-Schwartz and then the definition of $\|A\|$ we get

$$([I - A]x, x) = (x, x) - (Ax, x) \geq \|x\|^2 - \|Ax\| \, \|x\| \geq \|x\|^2 - \|A\|\|x\|^2 \geq 0$$

so $(I - A) \geq 0$ and applied to $-A$ gives $I + A \geq 0$ or $-I \leq A \leq I$.

Conversely, suppose that $-I \leq A \leq I$. Since $I - A \geq 0$ we know that $\mathrm{Spec}(A) \subset (-\infty, 1]$ and since $I + A \geq 0$ we have $\mathrm{Spec}(A) \subset (-1, \infty]$. So

$$\mathrm{Spec}(A) \subset [-1, 1]$$

so that the spectral radius of A is ≤ 1. But for self adjoint operators we have $\|A^2\| = \|A\|^2$ and hence the formula for the spectral radius gives $\|A\| \leq 1$. QED

An immediate corollary of the theorem is the following: Suppose that μ is a real number. Then $\|A - \mu I\| \leq \epsilon$ is equivalent to $(\mu - \epsilon)I \leq A \leq (\mu_+ \epsilon)I$. So one way of interpreting the spectral theorem

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda$$

is to say that for any doubly infinite sequence

$$\cdot < \lambda_{-2} < \lambda_{-1} < \lambda_0 < \lambda_1 < \lambda_2 < \cdots$$

with $\lambda_{-n} \to -\infty$ and $\lambda_n \to \infty$ there is a corresponding Hilbert space direct sum decomposition

$$H = \bigoplus H_i$$

invariant under A and such that the restriction of A to H_i satisfies

$$\lambda_i I \leq A_{|H_i} \leq \lambda_{i+1} I.$$

If $\mu_i := \frac{1}{2}(\lambda_i + \lambda_{i+1})$ then another way of writing the preceding inequality is

$$\|A_{H_i} - \mu_i I\| \leq \frac{1}{2}(\lambda_{i+1} - \lambda_i).$$

7.2 The point spectrum

We now let A denote an arbitrary (not necessarily bounded) self adjoint transformation. We say that λ belongs to the **point spectrum** of A if there exists an $x \in D(A)$ such that $x \neq 0$ and $Ax = \lambda x$. In other words if λ is an eigenvalue of A. Notice that eigenvectors corresponding to distinct eigenvalues are orthogonal: if $Ax = \lambda x$ and $Ay = \mu y$ then

$$\lambda(x, y) = (\lambda x, y) = (Ax, y) = (x, Ay) = (x, \mu y) = \mu(x, y)$$

implying that $(x, y) = 0$ if $\lambda \neq \mu$.

Also, the fact that a self-adjoint operator is closed implies that the space of eigenvectors corresponding to a fixed eigenvalue is a closed subspace of H. We let N_λ denote the space of eigenvectors corresponding to an eigenvalue λ.

We say that A has **pure point spectrum** if its eigenvectors span H, in other words if

$$H = \bigoplus N_{\lambda_i}$$

where the λ_i range over the set of eigenvalues of A. Suppose that this is the case. Then let

$$M_\lambda := \bigoplus_{\mu<\lambda} N_\mu$$

where this denotes the Hilbert space direct sum, i.e. the closure of the algebraic direct sum. Let E_λ denote projection onto M_λ. Then it is immediate that the E_λ satisfy (8)-(13) and that (14) holds with the interpretation given in the preceding section. We thus have a proof of the spectral theorem for operators with pure point spectrum.

7.3 Partition into pure types.

Now consider a general self-adjoint operator A, and let

$$H_1 := \bigoplus N_\lambda$$

(Hilbert space direct sum) and set

$$H_2 := H_1^\perp.$$

The space H_1 and hence the space H_2 are invariant under A in the sense that A maps $D(A) \cap H_1$ to H_1 and similarly for H_2.

We let P denote orthogonal projection onto H_1 so $I - P$ is orthogonal projection onto H_2. We claim that

$$P[D(A)] = D(A) \cap H_1 \quad \text{and} \quad (I - P)[D(A)] = D(A) \cap H_2. \tag{18}$$

Suppose that $x \in D(A)$. We must show that $Px \in D(A)$ for then $x = Px + (I - P)x$ is a decomposition of every element of $D(A)$ into a sum of elements of $D(A) \cap H_1$ and $D(A) \cap H_2$.

By definition, we can find an orthonormal basis of H_1 consisting of eigenvectors u_i of A, and then

$$Px = \sum a_i u_i \qquad a_i := (x, u_i).$$

The sum on the right is (in general) infinite. Let y denote any finite partial sum. Since eigenvectors belong to $D(A)$ we know that $y \in D(A)$. We have

$$(A[x-y], Ay) = ([x-y], A^2 y) = 0$$

since $x - y$ is orthogonal to all the eigenvectors occurring in the expression for y. We thus have

$$\|Ax\|^2 = \|A(x-y)\|^2 + \|Ay\|^2$$

We know that (as we let the number of terms in y increase) y converges to Px. From the preceding equality we know that $\sum |\lambda_i a_i|^2 \leq \|Ax\|^2 < \infty$ and hence that the Ay converge. Since A is closed, we conclude that $Px \in D(A)$ proving (18).

Let A_1 denote the operator A restricted to $P[D(A)] = D(A) \cap H_1$ with similar notation for A_2. We claim that A_1 is self adjoint (as is A_2). Clearly $D(A_1) := P(D(A))$ is dense in H_1, for if there were a vector $y \in H_1$ orthogonal to $D(A_1)$ it would be orthogonal to $D(A)$ in H which is impossible. Similarly $D(A_2) := D(A) \cap H_2$ is dense in H_2.

Now suppose that y_1 and z_1 are elements of H_1 such that

$$(A_1 x_1, y_1) = (x_1, z_1) \quad \forall\, x_1 \in D(A_1).$$

Since $A_1 x_1 = A x_1$ and $x_1 = x - x_2$ for some $x \in D(A)$, and since y_1 and z_1 are orthogonal to x_2, we can write the above equation as

$$(Ax, y_1) = (x, z_1) \quad \forall\, x \in D(A)$$

which implies that $y_1 \in D(A) \cap H_1 = D(A_1)$ and $A_1 y_1 = A y_1 = z_1$.

In other words, A_1 is self-adjoint. Similarly, so is A_2. We have thus proved

Theorem 11 *Let A be a self-adjoint transformation on a Hilbert space H. Then*

$$H = H_1 \oplus H_2$$

with self-adjoint transformations A_1 on H_1 having pure point spectrum and A_2 on H_2 having no point spectrum such that

$$D(A) = D(A_1) \oplus D(A_2)$$

and

$$A = A_1 \oplus A_2.$$

We have proved the spectral theorem for a self adjoint operator with pure point spectrum. Our proof of the full spectral theorem will be complete once we prove it for operators with no point spectrum.

7.4 Completion of the proof.

In this subsection we will assume that A is a self-adjoint operator with no point spectrum, i.e. no eigenvalues.

Let $\lambda < \mu$ be real numbers and let C be a closed piecewise smooth curve in the complex plane which is symmetrical about the real axis and cuts the real axis at non-zero angle at the two points λ and μ (only). Let $m > 0$ and $n > 0$ be positive integers, and let

$$K_{\lambda\mu}(m,n) := \frac{1}{2\pi i}\int_C (z-\lambda)^m(z-\mu)^n R_z dz. \tag{19}$$

In fact, we would like to be able to consider the above integral when $m = n = 0$, in which case it should give us a projection onto a subspace where $\lambda I \leq A \leq \mu I$. But unfortunately if λ or μ belong to $\mathrm{Spec}(A)$ the above integral need not converge with $m = n = 0$. However we do know that $\|R_z\| \leq (|\operatorname{Im} z|)^{-1}$ so that the blow up in the integrand at λ and μ is killed by $(z-\lambda)^m$ and $(\mu - z)^n$ since the curve makes non-zero angle with the real axis. Since the curve is symmetric about the real axis, the (bounded) operator $K_{\lambda\mu}(m,n)$ is self-adjoint. Furthermore, modifying the curve C to a curve C' lying inside C, again intersecting the real axis only at the points λ and μ and having these intersections at non-zero angles does not change the value: $K_{\lambda\mu}(m,n)$.

We will now prove a succession of facts about $K_{\lambda\mu}(m,n)$:

$$K_{\lambda\mu}(m,n)\cdot K_{\lambda\mu}(m',n') = K_{\lambda\mu}(m+m',n+n'). \tag{20}$$

Proof. Calculate the product using a curve C' for $K_{\lambda\mu}(m',n')$ as indicated above. Then use the functional equation for the resolvent and Cauchy's integral formula exactly as in the proof of Theorem 7: $(2\pi i)^2 K_{\lambda\mu}(m,n)\cdot K_{\lambda\mu}(m',n') =$

$$\int_C\int_{C'}(z-\lambda)^m(\mu-z)^n(w-\lambda)^{m'}(\mu-w)^{n'}\frac{1}{z-w}[R_w - R_z]dzdw$$

which we write as a sum of two integrals, the first giving

$$(2\pi i)^2 K_{\lambda\mu}(m+m',n+n')$$

and the second giving zero. QED

A similar argument (similar to the proof of Theorem 8) shows that

$$K_{\lambda\mu}(m,n)\cdot K_{\lambda'\mu'}(m',n') = 0 \ \text{ if } (\lambda,\mu)\cap(\lambda',\mu') = \emptyset. \tag{21}$$

Proposition 8 *There exists a bounded self-adjoint operator $L_{\lambda\mu}(m,n)$ such that*

$$L_{\lambda\mu}(m,n)^2 = K_{\lambda\mu}(m,n).$$

Proof. The function $z \mapsto (z-\lambda)^{m/2}(\mu-z)^{n/2}$ is defined and holomorphic on the complex plane with the closed intervals $(-\infty,\lambda]$ and $[\mu,\infty)$ removed. The integral

$$L_{\lambda\mu}(m,n) = \frac{1}{2\pi i}\int_C (z-\lambda)^{m/2}(\mu-z)^{n/2}R_z dz$$

is well defined since, if $m = 1$ or $n = 1$ the singularity is of the form $|\operatorname{Im} z|^{-\frac{1}{2}}$ at worst which is integrable. Then the proof of (20) applies to prove the proposition. QED

For each complex z we know that $R_z x \in D(A)$. Hence

$$(A - \lambda I)R_z x = (A - zI)R_z x + (z - \lambda)R_z x = x + (z - \lambda)R_z x.$$

By writing the integral defining $K_{\lambda\mu}(m, n)$ as a limit of approximating sums, we see that $(A - \lambda I)K_{\lambda\mu}(m, n)$ is defined and that it is given by the sum of two integrals, the first of which vanishes and the second gives $K_{\lambda\mu}(m + 1, n)$.

We have thus shown that $K_{\lambda\mu}(m, n)$ maps H into $D(A)$ and

$$(A - \lambda I)K_{\lambda\mu}(m, n) = K_{\lambda\mu}(m + 1, n). \tag{22}$$

Similarly

$$(\mu I - A)K_{\lambda\mu}(m, n) = K_{\lambda\mu}(m, n + 1). \tag{23}$$

We also have

$$\lambda(x, x) \leq (Ax, x) \leq \mu(x, x) \quad \text{for } x \in \text{ the image of } K_{\lambda\mu}(m, n). \tag{24}$$

Proof. We have

$$\begin{aligned}
([A - \lambda I]K_{\lambda\mu}(m, n)y, K_{\lambda\mu}(m, n)y) &= (K_{\lambda\mu}(m + 1, n)y, K_{\lambda\mu}(m, n)y) \\
&= (K_{\lambda\mu}(m, n)K_{\lambda\mu}(m + 1, n)y, y) \\
&= (K_{\lambda\mu}(2m + 1, 2n)y, y) \\
&= (L_{\lambda\mu}(2m + 1, 2n)^2 y, y) \\
&= (L_{\lambda\mu}(2m + 1, 2n)y, L_{\lambda\mu}(2m + 1, 2n)y) \geq 0. \\
&\Rightarrow (Ax, x) \geq \lambda(x, x) \quad \text{if} \\
x &= K_{\lambda\mu}(m, n)y.
\end{aligned}$$

A similar argument proves the second inequality in (24). QED

Thus if we define $M_{\lambda\mu}(m, n)$ to be the closure of the image of $K_{\lambda\mu}(m, n)$ we see that A is bounded when restricted to $M_{\lambda\mu}(m, n)$ and

$$\lambda I \leq A \leq \mu I$$

there.

We let $N_{\lambda\mu}(m, n)$ denote the kernel of $K_{\lambda\mu}(m, n)$ so that $M_{\lambda\mu}(m, n)$ and $N_{\lambda\mu}(m, n)$ are the orthogonal complements of one another since $K_{\lambda\mu}(m, n)$ is a bounded self-adjoint operator.

So far we have not made use of the assumption that A has no point spectrum. Here is where we will use this assumption: Since

$$(A - \lambda I)K_{\lambda\mu}(m, n) = K_{\lambda\mu}(m + 1, n)$$

we see that if $K_{\lambda\mu}(m + 1, n)x = 0$ we must have $(A - \lambda I)K_{\lambda\mu}(m, n)x = 0$ which, by our assumption implies that $K_{\lambda\mu}(m, n)x = 0$. In other words,

Proposition 9 *The space $N_{\lambda\mu}(m,n)$, and hence its orthogonal complement $M_{\lambda\mu}(m,n)$ is independent of m and n.*

We will denote the common space $M_{\lambda\mu}(m,n)$ by $M_{\lambda\mu}$. We have proved that A is a bounded operator when restricted to $M_{\lambda\mu}$ and satisfies

$$\lambda I \leq A \leq \mu I \quad \text{on } M_{\lambda\mu}$$

there.

We now claim that

$$\text{If } \lambda < \nu < \mu \text{ then } \quad M_{\lambda\mu} = M_{\lambda\nu} \oplus M_{\nu\mu}. \tag{25}$$

Proof. Let $C_{\lambda\mu}$ denote the rectangle of height one parallel to the real axis and cutting the real axis at the points λ and μ. Use similar notation to define the rectangles $C_{\lambda\nu}$ and $C_{\nu\mu}$. Consider the integrand

$$S_z := (z-\lambda)(z-\mu)(z-\nu)R_z$$

and let

$$T_{\lambda\mu} := \frac{1}{2\pi i}\int_{C_{\lambda\mu}} S_z dz$$

with similar notation for the integrals over the other two rectangles of the same integrand. Then clearly

$$T_{\lambda\mu} = T_{\lambda\nu} + T_{\nu\mu} \quad \text{and } T_{\lambda\nu} \cdot T_{\nu\mu} = 0. \tag{26}$$

Also, writing $zI - A = (z-\nu)I + (\nu I - A)$ we see that

$$(\nu I - A)K_{\lambda\mu}(1,1) = T_{\lambda\mu}$$

Since A has no point spectrum, the closure of the image of $T_{\lambda\mu}$ is the same as the closure of the image of $K_{\lambda\mu}(1,1)$, namely $M_{\lambda\mu}$. Thus (25) follows from (26).

If we now have a doubly infinite sequence as in our reformulation of the spectral theorem, and we set $M_i := M_{\lambda_i\lambda_{i+1}}$ we have proved the spectral theorem (in the no point spectrum case - and hence in the general case) if we show that

$$\bigoplus M_i = H.$$

In view of (25) it is enough to prove that the closure of the limit of M_{-rr} is all of H as $r \to \infty$, or, what amounts to the same thing, if y is perpendicular to all $K_{-rr}(1,1)x$ then y must be zero. But

$$(K_{-rr}(1,1)x, y) = (x, K_{-rr}(1,1)y).$$

So we must show that if $K_{-rr}y = 0$ for all r then $y = 0$. Now

$$K_{-rr} = \frac{1}{2\pi i}\int_C (z+r)(r-z)R_z dz = -\frac{1}{2\pi i}\int (z^2 - r^2)R_z$$

where we may take C to be the circle of radius r centered at the origin. We also have

$$1 = \frac{1}{2\pi i r^2}\int_C \frac{r^2 - z^2}{z} dz.$$

So

$$y = \frac{1}{2\pi i r^2}\int_C (r^2 - z^2)[z^{-1}I - R_z]dz \cdot y.$$

But $(zI - A)R_z = I$ so $-AR_z = I - zR_z$ or

$$z^{-1}I - R_z = -z^{-1}AR_z$$

so (pulling the A out from under the integral sign) we can write the above expression for y as

$$y = Ag_r \quad \text{where } g_r = \frac{1}{2\pi i r^2}\int_C (r^2 - z^2)z^{-1}R_z dz \cdot y.$$

On C we have $z = re^{i\theta}$ so $z^2 = r^2 e^{2i\theta} = r^2(\cos 2\theta + i \sin 2\theta)$ and hence

$$z^2 - r^2 = r^2(\cos 2\theta - 1 + i \sin 2\theta) = 2r^2(-\sin^2\theta + i \sin\theta\cos\theta).$$

Now $\|R_z\| \leq |r \sin\theta|^{-1}$ so we see that

$$\|(z^2 - r^2)R_z\| \leq 4r.$$

Since $|z^{-1}| = r^{-1}$ on C, we can bound $\|g_r\|$ by

$$\|g_r\| \leq (2\pi r^2)^{-1} \cdot r^{-1} \cdot 4r \cdot 2\pi r\|y\| = 4r^{-1}\|y\| \to 0$$

as $r \to \infty$. We have proved that $y = Ag_r$ and that $g_r \to 0$. Since A is closed (being self-adjoint) we conclude that if we defined $y_r \equiv y$ then $0 = \lim y_r$, i.e. that $y = 0$. This concludes Lorch's proof of the spectral theorem.

8 Appendix. The closed graph theorem.

Lurking in the background of our entire discussion is the closed graph theorem which says that if a closed linear transformation from one Banach space to another is everywhere defined, it is in fact bounded. We did not actually use this theorem, but its statement and proof by Banach greatly clarified the notion of a what an unbounded self-adjoint operator is, and explained the Hellinger Toeplitz theorem as I mentioned earlier. So here I will give the standard proof of this theorem (essentially a Baire category style argument) taken from Loomis.

In what follows X and Y will denote Banach spaces,

$$B_n := B_n(X) = \{x \in X;\ \|x\| \leq n\}$$

denotes the ball of radius n about the origin in X and

$$U_r = B_r(Y) = \{y \in Y:\ \|y\| \leq r\}$$

the ball of radius r about the origin in Y.

Lemma 1 *Let*

$$T : X \to Y$$

be a bounded (everywhere defined) linear transformation. If $T[B_1] \cap U_r$ is dense in U_r then

$$U_r subset T[B_1].$$

Proof. The set $T[B_1]$ is closed, so it will be enough to show that

$$U_{r(1-\delta)} subset T[B_1]$$

for any $\delta > 0$, or, what is the same thing, that

$$U_r subset \frac{1}{1-\delta} T[B_1] = T[B_{\frac{1}{1-\delta}}].$$

So fix $\delta > 0$. Let $z \in U_r$. Set $y_0 := 0$, and choose $y_1 \in T[B_1] \cap U_r$ such that $\|z - y_1\| < \delta r$. Since $\delta\,(T[B_1] \cap U_r)$ is dense in δU_r we can find $y_2 - y_1 \in \delta\,(T[B_1] \cap U_r)$ within distance $\delta^2 r$ of $z - y_1$ which implies that $\|y_2 - z\| < \delta^2 r$. Proceeding inductively we find a sequence $\{y_n \in Y$ such that

$$y_{n+1} - y_n \in \delta^n \,(T[B_1] \cap U_r)$$

and

$$\|y_{n+1} - z\|, \delta^{n+1} r.$$

We can thus find $x_n \in X$ such that

$$T(x_{n+1} = y_{n+1} - y_n \quad \text{and} \quad \|x_{n+1}\| < \delta^n.$$

If

$$x := \sum_1^\infty x_n$$

then $\|x\| < 1/(1-\delta)$ and $Tx = z$. QED

Lemma 2 *If $T[B_1]$ is dense in no ball of positive radius in Y, then $T[X]$ contains no ball of positive radius in Y.*

Proof. Under the hypotheses of the lemma, $T[B_n]$ is also dense in no ball of positive radius of Y. So given any ball $U \subset Y$, we can find a (closed) ball U_{r_1,y_1} of radius r_1 about y_1 such that $U_{r_1,y_1} \subset U$ and is disjoint from $T[B_1]$. By induction, we can find a nested sequence of balls $U_{r_n,y_n} subset U_{r_{n-1} y_{n-1}}$ such that U_{r_n,y_n} is disjoint form $T[B_n]$ and can also arrange that $r_n \to 0$. Choosing a point in each of these balls we get a Cauchy sequence which converges to a point $y \in U$ which lies in none of the $T[B_n]$, i.e. $y \not{i}n T[X]$. So $u \not\subset T[X]$. QED

Theorem 12 **[The bounded inverse theorem.]** *If $T : X \to Y$ is bounded and bijective, then T^{-1} is bounded.*

Proof. By Lemma 2, $T[B_1]$ is dense in some ball U_{r,y_1} and hence

$$T[B_1 + B_1] = T[B_1 - B_1]$$

is dense in a ball of radius r about the origin. Since $B_1 + B_1 \subset B_2$ so $T[B_2] \cap U_r$ is dense in U_r. By Lemma 1, this implies that

$$T[B_2] \supset U_r$$

i.e. that

$$T^{-1}[U_r] \subset B_2$$

which says that

$$\|T^{-1}\| \leq \frac{2}{r}.$$

QED

Theorem 13 *If $T : X \to Y$ is defined on all of X and is such that* graph(T) *is a closed subspace of $X \oplus Y$, then T is bounded.*

Proof. Let $\Gamma \subset X \oplus Y$ denote the graph of T. By assumption, it is a closed subspace of the Banach space $X \oplus Y$ under the norm $\|\{x, y\}\| = \|x\| + \|y\|$. So Γ is a Banach space and the projection

$$\Gamma \to X, \quad \{x, y\} \mapsto x$$

is bijective by the definition of a graph, and has norm ≤ 1. So its inverse is bounded. Similarly the projection onto the second factor is bounded. So the composite map

$$X \to Y \quad x \mapsto \{x, y\} \mapsto y = Tx$$

is bounded. QED