

Functional Analysis II (Math 921)

Lecture Notes for Fall '08

Jeff Schenker

MICHIGAN STATE UNIVERSITY

E-mail address: `jeffrey@math.msu.edu`

Contents

Part 1. Spectral Theory in Banach Algebras

- Lecture 1. Spectral Theory– an Introduction
- Lecture 2. The spectrum in Banach algebras
- Lecture 3. Spectral radius and the analytic functional calculus
- Lecture 4. Spectral mapping theorem and Riesz Projections
- Lecture 5. Dependence of the spectrum on the algebra
- Lecture 6. Commutative Banach Algebras
- Lecture 7. Spectral theory in commutative Banach Algebras
- Lecture 8. C^* algebras

Part 2. Spectral Theory for operators on a Hilbert space

- Lecture 9. Self-adjoint operators
- Lecture 10. Functional calculus and polar decomposition
- Lecture 11. Spectral resolution
- Lecture 12. Representation of a self-adjoint operator as a multiplication operator
- Lecture 13. Spectrum, Spectral Measures and Spectral Multiplicities
- Lecture 14. Jacobi matrix representations and orthogonal polynomials
- Lecture 15. Generalized eigenvectors for Jacobi matrices
- Lecture 16. Unbounded operators
- Lecture 17. Functional calculus for self-adjoint operators
- Lecture 18. Spectral representation of self-adjoint operators
- Lecture 19. Unitary operators and von Neumann's proof the spectral theorem; Positive operators and the polar decomposition
- Lecture 20. Examples of self-adjoint extensions
- Lecture 21. Theory of self-adjoint extensions
- Lecture 22. Self-adjoint extensions (II) and the Friedrichs extension

Part 3. Semigroups

Lecture 23. Continuous and strongly continuous semigroup

Lecture 24. Unbounded operators in a Banach space

Lecture 25. The Hille-Yosida theorem

Lecture 26. Unitary semigroups and self-adjoint semigroups

Part 4. Perturbation Theory

Lecture 27. Rellich's Theorem

Lecture 28. Perturbation of eigenvalues

Part 1

Spectral Theory in Banach Algebras

LECTURE 1

Spectral Theory— an Introduction

In this course we are going to focus on spectral theory for linear operators. The goal of spectral theory is to understand at a detailed level how a linear operator acts on the vector space on which it is defined. One key reason for doing this is to make sense of solving equations, like

$$\partial_t \psi_t = A \psi_t, \quad (1.1)$$

where ψ is a vector in a Banach space, say, and A is a linear operator. If A is bounded — recall that this means $\|A\psi\| \leq C \|\psi\|$ for all ψ in the Banach space — then we can solve this equation somewhat formally using the series for the exponential:

$$\psi_t = \psi_0 + \sum_{n=1}^{\infty} \frac{t^n}{n!} A^n \psi_0. \quad (1.2)$$

EXERCISE 1. Show, if A is bounded then (1.2) defines a solution to (1.1), with $\partial_t \psi_t$ defined as derivative in norm. That is

$$\lim_{h \rightarrow 0} \left\| \frac{1}{h} (\psi_{t+h} - \psi_t) - \partial_t \psi_t \right\| = 0.$$

This may not be a particularly efficient procedure, and it can fail completely for some unbounded operators of interest.

For instance, the class of functions $\psi_0 \in L^2(\mathbb{R})$ for which we can solve the Schrödinger equation

$$i \partial_t \psi_t(x) = -\partial_x^2 \psi_t(x) \quad (1.3)$$

in this way is quite limited. Nonetheless, (1.3) can be solved for any initial $\psi_0 \in L^2(\mathbb{R})$ using a Fourier transform

$$\widehat{\psi}(k) = \int_{-\infty}^{\infty} e^{2\pi i k x} \psi(x) dx. \quad (1.4)$$

As is well known $\psi \mapsto \widehat{\psi}$ defines a unitary (norm preserving) map from $L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$. (Strictly speaking (1.4) is defined only on $L^1(\mathbb{R})$. It has to be extended to $L^2(\mathbb{R})$ by taking limits.) Furthermore one can show, using integration by parts, that ψ_t satisfies (1.3) if and only

$$i \partial_t \widehat{\psi}_t(k) = 4\pi^2 k^2 \widehat{\psi}_t(k),$$

which is easily solved to give

$$\widehat{\psi}_t(k) = e^{-it4\pi^2 k^2} \widehat{\psi}_0(k).$$

It is now simply a matter of inverting the Fourier transform via

$$\psi(x) = \int_{-\infty}^{\infty} e^{-2\pi i m k x} \widehat{\psi}(k) dk$$

to find the solution to (1.3).

What we have done here is used the Fourier transform to diagonalize the linear operator $-\partial_x^2$. When we put the operator in diagonal form, namely multiplication by $4\pi^2 k^2$, it becomes easy to solve the differential equation (1.1). This example is the best sort of thing that can happen and is by no means typical.

Even if we stick to finite dimensional spaces, you might recall that not every matrix can be diagonalized. It may be useful before going on to recall the spectral theory of matrices.

A first thing you can try to do is to put a matrix in triangular form. That is given $A = (a_{i,j})_{i,j=1}^n$ we look for an invertible matrix S such that $B = SAS^{-1} = (b_{i,j})_{i,j=1}^n$. This can be done, of course, as follows: Since $\det(A - \lambda I)$ is a polynomial, it has a root. This root is an eigenvalue. So there is an eigenvector, u_1 with eigenvalues λ_1 . Now let $u_2^{(1)}, \dots, u_n^{(n)}$ be vectors which together with u_1 form a basis. The matrix of A in this basis is of the form

$$A \sim \begin{pmatrix} \lambda_1 & a_{1,2}^{(1)} & \cdots & a_{1,n}^{(1)} \\ 0 & a_{2,2}^{(1)} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n}^{(n)} \end{pmatrix},$$

Now repeat the argument with the lower right hand $(n-1) \times (n-1)$ block. The result, after n steps, is an upper triangular matrix.

Unfortunately, there is no analogue of this algorithm for a general bounded operator. There is a triangular form for compact operators, however, which we will discuss.

When it comes to diagonalizing a matrix, assumptions are needed. You may recall the Jordan form, which states that every matrix can be transformed, via a similarity transformation, into a matrix with zero entries except for the diagonal and first super diagonal, and furthermore only ones and zeros on the super diagonal. However, the number of ones in each block on the super diagonal is a characteristic of the matrix.

If we want a criteria for diagonalizability, one of the most useful requires a scalar product.

THEOREM 1.1. *If $\langle \cdot, \cdot \rangle$ denotes any inner product on $\mathbb{C}^n$ and A is a matrix such that*

$$AA^\dagger - A^\dagger A = 0, \tag{1.5}$$

— where $A^\dagger$ is the adjoint of A , $\langle u, Av \rangle = \langle A^\dagger u, v \rangle$ —, *then A is diagonalizable via a unitary transformation. In other words, there is an orthonormal basis of eigenvectors.*

REMARK. A matrix (or operator) which satisfies (1.5) is called normal.

PROOF. Let u_1 be an eigenvector with eigenvalues λ_1 . It suffices to show

$$\{u_1\}^\perp = \{u : \langle u, u_1 \rangle = 0\}$$

is an invariant subspace for A and $A^\dagger$, for then we can complete the proof by induction.

Suppose $u \in \{u_1\}^\perp$. Then

$$\langle A^\dagger u, u_1 \rangle = \langle u, Au_1 \rangle = \lambda_1 \langle u, u_1 \rangle = 0.$$

That is, $A^\dagger : \{u_1\}^\perp \rightarrow \{u_1\}^\perp$. Now note that

$$\begin{aligned} 0 &= \langle (A - \lambda_1 I)u_1, (A - \lambda_1 I)u_1 \rangle = \langle (A^\dagger - \lambda_1^* I)(A - \lambda_1 I)u_1, u_1 \rangle \\ &= \langle (A - \lambda_1 I)(A^\dagger - \lambda_1^* I)u_1, u_1 \rangle = \langle (A^\dagger - \lambda_1^* I)u_1, (A^\dagger - \lambda_1^* I)u_1 \rangle, \end{aligned}$$

so $\|(A^\dagger - \lambda^* I)u_1\| = 0$. In other words, u_1 is also an eigenvector of $A^\dagger$ with eigenvalue λ^* . Thus we can reverse the argument above and show that $A : \{u_1\}^\perp \rightarrow \{u_1\}^\perp$. $\square$

This theorem — diagonalizability of normal operators — has a far reaching generalization, namely the spectral theorem for normal operators, which we will see in due course.

The notion of normal depends on the scalar product. In the finite dimensional context, we have many choices for a scalar product and one can easily show that a matrix is diagonalizable if and only if it is normal in some inner product. (We just saw the if. For the only if take the inner product in which the eigenvectors are orthonormal.) This question becomes trickier in the infinite dimensional context.

1. Outline of the first part of the course

- (1) We are going to start in an abstract setting and derive some general results on spectral theory of elements of Normed algebras (also called Banach algebras).
- (2) We will then look at the implication of these results in the Hilbert space setting.
- (3) Next we specialize to the case of normal operators in a Hilbert space and derive the spectral theorem.
- (4) Then we will turn to unbounded operators, which are exceedingly important in applications, and require a good deal more care.

We'll see where to go from there. I would like to cover, also, semi-group theory, more results on compact operators — things like trace class, and the definition of determinant and trace. Then I would like to look at more details of spectral theory, spectral measures, scattering theory and perturbation theory. We'll see what time allows.

LECTURE 2

The spectrum in Banach algebras

Reading: §11.4 and Ch. 17 in Lax

A Banach algebra is a Banach space (completed normed space) $\mathcal{A}$ on which we have defined an associative product so that $\mathcal{A}$ is an algebra and such that

$$\|AB\| \leq \|A\| \|B\|, \quad \|cA\| = |c| \|A\|.$$

The motivating example of a Banach algebra is the space $\mathcal{L}(X)$ of bounded linear maps from a Banach space X into itself. It turns out that a good deal of the spectral theory of linear maps can be carried out in the more general context of Banach algebras. A Banach algebra may or may not have a unit I , which is an element such that $IA = AI = A$ for all A .

Here are some basic facts about Banach algebras. For the proofs see Lax (or work it out!):

- (1) If $\mathcal{A}$ has a unit, then the unit is unique.
- (2) Any Banach algebra is a closed sub-algebra of an algebra with a unit. (If $\mathcal{A}$ has a unit there is nothing to prove. If not, consider the space $\mathcal{A} \oplus \mathbb{C}$ with product

$$(A, z)(B, w) = (AB + zB + wA, zw)$$

and norm

$$\|(A, z)\| = \|A\| + |z|.$$

- (3) An element $A \in \mathcal{A}$ is invertible if there exists $B \in \mathcal{A}$ such that $BA = AB = I$.
- (4) It can happen that A has either a left inverse ($BA = I$) or a right inverse ($AB = I$) but is not invertible.

EXERCISE 2. Find an example of an operator in $\mathcal{L}(\ell^2(\mathbb{N}))$ that has a left inverse but no right inverse. Can you find one in $\mathcal{L}(\mathbb{C}^n)$? Why?

- (5) If A has a left inverse B and a right inverse C then $B = C$ and A is invertible.

THEOREM 2.1. *The set of invertible elements in $\mathcal{A}$ is open. Specifically, if A is invertible then $A + K$ is invertible provided $\|K\| < 1/\|A^{-1}\|$.*

EXERCISE. Using the geometric series to prove this theorem. (Or if you feel lazy, read the proof in Lax.)

DEFINITION 2.1. The resolvent set of $\rho(A)$ of A is the set of $\zeta \in \mathbb{C}$ such that

$$\zeta I - A$$

is invertible. The spectrum $\sigma(A)$ of A is the complement of the resolvent set. The resolvent of A is the map $R : \rho(A) \rightarrow \mathcal{A}$ given by

$$R(\zeta) = (\zeta I - A)^{-1}.$$

It turns out that $R(\zeta)$ is an analytic function. To make sense of this, we should first consider what it means for a Banach space valued function to be analytic.

1. Interlude: Analytic functions

Let X be a Banach space.

DEFINITION 2.2. A function $f : U \rightarrow X$, with $U \subset \mathbb{C}$ and open set, is strongly analytic (or just analytic) if

$$\lim_{h \rightarrow 0} \frac{1}{h} [f(\zeta + h) - f(\zeta)]$$

exists as a norm limit for every $\zeta \in U$, in which case the limit is denoted $f'(\zeta)$ or $\frac{d}{d\zeta}f(\zeta)$.

If f is an (Banach space valued) analytic function, the following familiar facts from complex analysis hold:

- (1) If f is analytic so is f' .
- (2) f is analytic if and only if
 - (a) f has a convergent power series expansion at each point in its domain.
 - (b) f is continuous and $f(\zeta) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) \frac{1}{z-\zeta} dz$ for any rectifiable closed curve $\mathcal{C}$ that can be contracted to a point in U and with winding number 1 around ζ .
(The integral can be taken as a Riemann integral since f is continuous.)

EXERCISE 3. Prove these facts. While you are at it, verify that the Riemann integral can be defined for norm continuous functions taking values in a Banach space.

You might also define what looks like a weaker notion of analyticity:

DEFINITION 2.3. f is weakly analytic in U if for every $\ell \in X' =$ the dual of X , $\ell(f(\zeta))$ is a (scalar) analytic function.

Clearly if f is strongly analytic it is weakly analytic. (Why?) What is surprising is the following

THEOREM 2.2. *If f is weakly analytic it is strongly analytic.*

PROOF. By the Cauchy integral formula

$$\ell(f(\zeta)) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\ell(f(z))}{z - \zeta} dz$$

for a suitably chosen curve $\mathcal{C}$. This formula holds if ζ is moved a little bit, so

$$\begin{aligned} & \ell \left(\frac{f(\zeta + h) - f(\zeta)}{h} - \frac{f(\zeta + k) - f(\zeta)}{k} \right) \\ &= \frac{1}{2\pi i} \int_{\mathcal{C}} \left[\frac{1}{h} \left(\frac{1}{z - h - \zeta} - \frac{1}{z - \zeta} \right) - \frac{1}{k} \left(\frac{1}{z - k - \zeta} - \frac{1}{z - \zeta} \right) \right] \ell(f(z)) dz \\ &= \frac{1}{2\pi i} \int_{\mathcal{C}} \left[\frac{1}{z - h - \zeta} - \frac{1}{z - k - \zeta} \right] \frac{\ell(f(z))}{z - \zeta} dz \\ &= \frac{(h - k)}{2\pi i} \int_{\mathcal{C}} \frac{1}{(z - h - \zeta)(z - k - \zeta)} \frac{\ell(f(z))}{z - \zeta} dz. \end{aligned}$$

It follows that

$$\left| \ell \left(\frac{1}{h - k} \left[\frac{f(\zeta + h) - f(\zeta)}{h} - \frac{f(\zeta + k) - f(\zeta)}{k} \right] \right) \right| \leq M(\ell) < \infty.$$

uniformly for all h, k sufficiently close to 0. By the Principle of Uniform Boundedness, there is a constant $C < \infty$ such that

$$\left\| \frac{f(\zeta + h) - f(\zeta)}{h} - \frac{f(\zeta + k) - f(\zeta)}{k} \right\| \leq C|h - k|.$$

Thus the limit defining $f'(\zeta)$ exists and f is analytic. $\square$

2. Back to the spectrum

THEOREM 2.3. *Let $\mathcal{A}$ be a Banach algebra. For any $A \in \mathcal{A}$ the resolvent $R(\zeta)$ is analytic on the resolvent set and the spectrum $\sigma(A)$ is a non-empty, compact subset of $\{\zeta \leq \|A\|\}$.*

PROOF. The Neumann series shows $R(\zeta)$ has a convergent power series at each point:

$$R(\zeta + h) = ((\zeta + h)I - A)^{-1} = \sum_{n=0}^{\infty} h^n (\zeta I - Z)^{n+1}$$

for small enough h . Analyticity follows.

Furthermore, for $\zeta > \|A\|$ we have

$$R(\zeta) = \sum_{n=0}^{\infty} A^n \frac{1}{\zeta^{n+1}}, \quad (2.1)$$

so $\sigma(A) \subset \{\zeta \leq \|A\|\}$. Since $\rho(A)$ is open it follows that $\sigma(A)$ is compact.

Integrating (2.1) around a large circle, we obtain

$$\int_{|\zeta|=r} R(\zeta) d\zeta = 2\pi i I$$

for $r > \|A\|$. Suppose $\sigma(A)$ were empty. Then we could contract the circle down to a point and would obtain 0. Since the result is not 0, $\sigma(A)$ is not empty. $\square$

LECTURE 3

Spectral radius and the analytic functional calculus

1. Spectral radius

Last time we saw that

$$\sigma(A) \subset \{z \leq \|A\|\}.$$

The question comes up as to whether this estimate is sharp. In other words if we let the spectral radius of A be

$$\text{sp-rad}(A) = \max\{|z| : z \in \sigma(A)\},$$

then can it happen that

$$\text{sp-rad}(A) < \|A\|?$$

A moment's thought shows this can happen for 2×2 matrices:

EXERCISE 4. Find a 2×2 matrix M with $\|M\| = 1$ and $\text{sp-rad } M = 0$.

However, we do have the following

THEOREM 3.1. $\text{sp-rad } A = \lim_{n \rightarrow \infty} \|A^n\|^{1/n}$.

PROOF. Consider the Laurent expansion of the resolvent around ∞ :

$$R(\zeta) = \sum_{n=0}^{\infty} \zeta^{-n-1} A^n. \tag{3.1}$$

I claim that this converges whenever $|\zeta| > \|A^k\|^{\frac{1}{k}}$. Indeed we can write

$$R(\zeta) = \left[\sum_{m=0}^{k-1} \zeta^{-m-1} A^m \right] \left[\sum_{n=0}^{\infty} \zeta^{-nk} A^k \right].$$

The first factor is a finite sum and the second converges if the above condition holds. Thus

$$\text{sp-rad}(A) \leq \liminf_{k \rightarrow \infty} \|A^k\|^{\frac{1}{k}}.$$

On the other hand if we let $\mathcal{C} = \{|z| = \text{sp-rad}(A) + \delta\}$ then $\mathcal{C}$ is a curve in the resolvent set which winds once around the spectrum. Integrating on this curve and using (3.1) we find

$$\frac{1}{2\pi i} \int_{\mathcal{C}} \zeta^n R(\zeta) d\zeta = A^n.$$

Thus

$$\|A^n\| \leq \left[\sup_{\zeta \in \mathcal{C}} \|R(\zeta)\| \right] (\text{sp-rad}(A) + \delta)^{n+1}.$$

Taking the n^{th} root and sending n to infinity, we get

$$\limsup_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}} \leq \text{sp-rad}(A) + \delta.$$

(Why is $\sup_{\zeta \in \mathcal{C}} \|R(\zeta)\| < \infty$?)

Since δ was arbitrary, we have

$$\limsup \|A^n\|^{\frac{1}{n}} \leq \text{sp-rad}(A) \leq \liminf \|A^n\|^{\frac{1}{n}}$$

and the result follows. $\square$

2. Functional calculus

The Cauchy integral formula

$$f(\zeta) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f(z)}{z - \zeta} dz$$

for scalar analytic functions suggests a way of defining $f(A)$ for A in a Banach algebra.

DEFINITION 3.1. Given an analytic function f , defined in a neighborhood U of $\sigma(A)$, we define

$$f(A) := \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) R(z) dz, \quad (3.2)$$

with R the resolvent of A and $\mathcal{C}$ any curve that has winding number 1 around the spectrum of A and winding number 0 around any point in U^c .

EXERCISE 5. Verify that this definition does not depend on the choice of curve $\mathcal{C}$.

That's that. We have defined $f(A)$ for any analytic function. For instance

$$A^2 = \frac{1}{2\pi i} \int_{\mathcal{C}} z^2 R(z) dz.$$

WAIT!!!!!! We can't define A^2 it's already defined! We need to check something. Does this definition make sense? Yes it does.

THEOREM 3.2. Let $f(A)$ be as defined above. Then

- (1) For any polynomial p , $p(A)$ (evaluated by algebra) is equal to the r.h.s. of (3.2).
- (2) More generally if f has a power series $f(z) = \sum_n a_n (z - z_0)^n$ convergent in a disk $\{|z - z_0| < r\}$ which contains $\sigma(A)$ then

$$\sum_{n=0}^{\infty} a_n (A - z_0 I)^n$$

is norm convergent and agrees with the r.h.s. of (3.2).

- (3) If f and g are two analytic functions defined in a neighborhood of $\sigma(A)$ then

$$f(A)g(A) = [fg](A).$$

EXERCISE 6. Prove (1) and (2).

PROOF. To prove (3) we will use the following

LEMMA 3.3 (Resolvent identity). Let $z, w \in \rho(A)$. Then

$$R(z) - R(w) = (w - z)R(z)R(w).$$

EXERCISE 7. Prove the resolvent identity.

Now let f, g be analytic in a neighborhood of $\sigma(A)$:

$$f(A) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(z) R(z) dz, \quad g(A) = \frac{1}{2\pi i} \int_{\mathcal{D}} g(z) R(z) dz.$$

Without loss of generality, assume $\mathcal{C}$ lies inside $\mathcal{D}$ — so the winding number of $\mathcal{D}$ around any point $z \in \mathcal{C}$ is one. Then

$$f(A)g(A) = \frac{1}{2\pi i} \int_{\mathcal{C}} \int_{\mathcal{D}} f(z)g(w)R(z)R(w)dw dz = \frac{1}{2\pi i} \int_{\mathcal{C}} \int_{\mathcal{D}} \frac{f(z)g(w)}{w-z} [R(z) - R(w)] dw dz.$$

Let's compute each term separately:

$$\begin{aligned} \frac{1}{2\pi i} \int_{\mathcal{C}} \int_{\mathcal{D}} \frac{f(z)g(w)}{w-z} R(z) dw dz &= \frac{1}{2\pi i} \int_{\mathcal{C}} f(z)g(z)R(z) dz = [fg](A), \\ \frac{1}{2\pi i} \int_{\mathcal{C}} \int_{\mathcal{D}} \frac{f(z)g(w)}{w-z} R(w) dw dz &= \frac{1}{2\pi i} \int_{\mathcal{D}} \left[\int_{\mathcal{C}} \frac{f(z)}{w-z} dz \right] g(w)R(w) dw = 0. \end{aligned}$$

□

The map $A \mapsto f(A)$ is called the functional calculus. Part (2) of the functional calculus shows that $f \mapsto f(A)$ is an algebraic homomorphism of the algebra of functions analytic in a neighborhood of $\sigma(A)$ into the Banach algebra $\mathcal{A}$. Next we will consider some of the analytic properties of this homomorphism.

LECTURE 4

Spectral mapping theorem and Riesz Projections

THEOREM 4.1. *Let $\mathcal{A}$ be a Banach Algebra, $A \in \mathcal{A}$, and f analytic in a neighborhood of $\sigma(A)$.*

- (1) *(The spectral mapping theorem): $\sigma(f(A)) = f(\sigma(A))$*
- (2) *If g is analytic in a neighborhood of $\sigma(f(A))$ then*

$$g(f(A)) = [g \circ f](A).$$

PROOF. To show (1) we need to show that $\zeta I - f(A)$ is invertible if and only if $\zeta \notin f(\sigma(A))$. If $\zeta \notin f(\sigma(A))$ then $h(z) = (\zeta - f(z))^{-1}$ is analytic in a neighborhood of $\sigma(A)$. But then

$$h(A)(\zeta I - f(A)) = I$$

by the multiplicative property of the functional calculus. On the other hand, if $\zeta \in f(\sigma(A))$, say $\zeta = f(w)$ with $w \in \sigma(A)$. Let

$$k(z) = \frac{f(w) - f(z)}{w - z},$$

so k is analytic in a neighborhood of $\sigma(A)$, and

$$k(A)(wI - A) = (wI - A)k(A) = \zeta I - f(A).$$

Suppose $(\zeta I - f(A))$ were invertible, then we would have

$$(\zeta I - f(A))^{-1}k(A)(wI - A) = (wI - A)k(A)(\zeta I - f(A))^{-1} = I,$$

which would imply that $w \notin \sigma(A)$, a contradiction.

To show (2), since $\sigma(f(A)) = f(\sigma(A))$ we have

$$g(f(A)) = \frac{1}{2\pi i} \oint_{\mathcal{D}} (\zeta I - f(A))^{-1} g(\zeta) d\zeta$$

with $\mathcal{D}$ a suitable contour. But

$$(\zeta I - f(A))^{-1} = \frac{1}{2\pi i} \oint_{\mathcal{C}} (zI - A)^{-1} \frac{1}{\zeta - f(z)} dz,$$

so

$$\begin{aligned} g(f(A)) &= \frac{1}{(2\pi i)^2} \oint_{\mathcal{D} \times \mathcal{C}} (zI - A)^{-1} \frac{1}{\zeta - f(z)} g(\zeta) dz d\zeta \\ &= \frac{1}{2\pi i} \oint_{\mathcal{C}} (zI - A)^{-1} g(f(z)) dz = [g \circ f](A). \quad \square \end{aligned}$$

The functional calculus $f \mapsto f(A)$ is often known as the “Riesz functional calculus” to distinguish it from the functional calculus we will develop later for self-adjoint operators, which will allow the evaluation of $f(A)$ for measurable functions.

One of the key conclusions of the spectral mapping theorem is the association of projections to each component of the spectrum of A .

DEFINITION 4.1. A projection P in $\mathcal{A}$ is any element of $\mathcal{A}$ which satisfies (1) $P^2 = P$ and (2) $P \neq 0$.

PROPOSITION 4.2. If P is a projection, then $P\mathcal{A}P = \{PAP : A \in \mathcal{A}\}$ is a Banach algebra with unit P . If $P \neq I$ then $\sigma(P) = \{0, 1\}$.

PROOF. Since $(PAP)(PBP) = P(APB)P$,

$$\|PAPPBP\| \leq \|PAP\| \|PBP\| \quad \text{and} \quad P(PAP) = (PAP)P = PAP,$$

it follows that $P\mathcal{A}P$ is a Banach algebra with unit P .

To see that $\sigma(P) = \{0, 1\}$, first note that

$$P(I - P) = (I - P)P = P - P = 0.$$

Thus neither P nor $(I - P)$ can be invertible. Since $P, (I - P) \neq 0$, it follows that $\{0, 1\} \subset \sigma(P)$.

It remains to show that any $\zeta \in \mathbb{C} \setminus \{0, 1\}$ is in the resolvent set. For this purpose, consider the Laurent series for the resolvent

$$(\zeta I - P)^{-1} = \sum_n \zeta^{-n-1} P^n,$$

convergent for $|\zeta| > \text{sp-rad}(P)$. Since $P^n = P$, $n \geq 1$, we may sum the series to get

$$R(\zeta) := (\zeta I - P)^{-1} = \frac{1}{\zeta}(I - P) + \sum_n \zeta^{-n-1} P = \frac{1}{\zeta}(I - P) + \frac{1}{\zeta - 1}P, \quad (4.1)$$

which is well defined for all $\zeta \in \mathbb{C} \setminus \{0, 1\}$.

EXERCISE 8. Check that the r.h.s. of (4.1) is equal to $(\zeta I - P)^{-1}$ for $\zeta \neq 0, 1$. □

Now suppose $A \in \mathcal{A}$ and

$$\sigma(A) = \cup_{j=1}^N \sigma_j$$

with σ_j disjoint. Then we can define

$$P_j = \frac{1}{2\pi i} \oint_{\mathcal{C}_j} (\zeta I - A)^{-1} d\zeta$$

with $\mathcal{C}_j$ any contour that winds once around σ_j and zero times around σ_i , $i \neq j$.

THEOREM 4.3.

- (1) P_j are projections
- (2) $P_j P_i = 0$ for $i \neq j$
- (3) $\sum_j P_j = I$.
- (4) The spectrum of $P_j A P_j = A P_j = P_j A$, as an element of the algebra $P_j \mathcal{A} P_j$, is

$$\sigma_{P_j \mathcal{A} P_j}(P_j A P_j) = \sigma_j.$$

REMARK. The spectrum of $P_j A P_j$ in $\mathcal{A}$ is $\sigma_j \cup \{0\}$.

PROOF. Note that $P_j = f_j(A)$ with f_j an analytic function that is 1 in a neighborhood of σ_j and 0 in a neighborhood of σ_i for $i \neq j$. Thus (1), (2) and (3) follow from the functional calculus. □

The projections P_j are known as “Riesz projections.” For matrices, they give the projection onto generalized eigenspaces. To see this, let us compute an example. Consider the matrix

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

The spectrum of A is $\{-1, +1\}$. If we write down the resolvent

$$(\zeta I - A)^{-1} = \frac{1}{\zeta^2 - 1} \begin{pmatrix} \zeta & 1 \\ 1 & \zeta \end{pmatrix},$$

then we may compute

$$P_{\pm} = \frac{1}{2\pi i} \oint_{z=\pm 1 + e^{i\theta}} \frac{1}{\zeta^2 - 1} \begin{pmatrix} \zeta & 1 \\ 1 & \zeta \end{pmatrix} d\zeta = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}.$$

EXERCISE 9. Verify that

$$P_{\pm}^2 = I, \quad AP_{\pm} = P_{\pm}A = \pm P_{\pm}.$$

More generally, the matrix may have non-trivial blocks in its Jordan form.

EXERCISE 10. Compute the resolvent and Riesz projections for

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

EXERCISE 11. Show that

$$f \left(\underbrace{\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ & \lambda & 1 & & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}}_{n \times n} \right) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ & f(\lambda) & f'(\lambda) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2}f''(\lambda) \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}$$

In infinite dimensions the Riesz projections may not be related to generalized eigenvectors. For instance the shift operator

$$S(a_0, a_1, \cdots) = (a_0, a_1, \cdots)$$

on ℓ^2 has spectrum

$$\sigma(S) = \{|z| \leq 1\}.$$

Thus S has only one Riesz projection — the identity map.

EXERCISE 12. S corresponds to the infinite matrix

$$S \sim \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \end{pmatrix}.$$

Show that, if f is analytic in a neighborhood of $\{|z| \leq 1\}$ then $f(S)$ corresponds to the infinite matrix

$$S \sim \begin{pmatrix} f(0) & f'(0) & \frac{1}{2}f''(0) & \cdots & \frac{1}{n!}f^{(n)}(0) & \cdots \\ & f(0) & f'(0) & \frac{1}{2}f''(0) & \cdots & \frac{1}{n!}f^{(n)}(0) & \cdots \\ & & \ddots & \ddots & \ddots & & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \ddots & \ddots \\ & & & & & & \ddots \end{pmatrix}.$$

In other words, if $\mathbf{a} = (a_0, a_1, a_2, \dots)$ then

$$[f(S)\mathbf{a}]_j = j^{th} \text{ entry of } f(S)\mathbf{a} = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) a_{j+n}.$$

LECTURE 5

Dependence of the spectrum on the algebra

Source: Some material in this lecture is taken from Chapter VII of John B. Conway, “A course in functional analysis,” 2nd ed., Springer 1990.

Recall that we defined the Riesz projections

$$P_j = \frac{1}{2\pi i} \oint_{\mathcal{C}_j} (\zeta I - A)^{-1} d\zeta$$

for $A \in \mathcal{A}$ with spectrum that can be written as a disjoint union

$$\sigma(A) = \cup_{j=1}^N \sigma_j.$$

I claimed that

- (1) $P_j A = A P_j = P_j A P_j$.
- (2) $\sigma_{\mathcal{A}_j}(P_j A P_j) = \sigma_j$, where $\sigma_{\mathcal{A}_j}$ denotes the spectrum in the Banach algebra $\mathcal{A}_j = P_j \mathcal{A} P_j$.

To see (1) note that

$$(\zeta I - A)^{-1} A = A (\zeta I - A)^{-1},$$

as one can see explicitly by adding and subtracting ζI from A . Or use the multiplicative property of the functional calculus. This gives the first equality of (1). The second follows since $P_j^2 = P_j$.

To see (2), we have to show that $(zI - P_j A P_j)$ is invertible in $\mathcal{A}_j$ if and only if $z \notin \sigma_j$. To this end note that

$$P_j A P_j = P_j A = \frac{1}{2\pi i} \int_{\mathcal{C}_j} \zeta (\zeta I - A)^{-1} d\zeta.$$

If $z \notin \sigma_j$ then $\zeta \mapsto (z - \zeta)^{-1}$ is analytic in a neighborhood of σ_j and by the functional calculus

$$\left[\frac{1}{2\pi i} \int_{\mathcal{C}_j} \frac{1}{z - \zeta} (\zeta I - A)^{-1} d\zeta \right] P_j A P_j = P_j A P_j \left[\frac{1}{2\pi i} \int_{\mathcal{C}_j} \frac{1}{z - \zeta} (\zeta I - A)^{-1} d\zeta \right] = P_j,$$

so $\zeta \notin \sigma_{\mathcal{A}_j}(P_j A P_j)$. On the other hand if $\zeta \in \sigma_j$ and there were B such that $(\zeta I - P_j A P_j) P_j B P_j = P_j B P_j (\zeta I - P_j A P_j) = P_j$ then

$$B + \sum_{i \neq j} \frac{1}{2\pi i} \int_{\mathcal{C}_i} \frac{1}{z - \zeta} (\zeta I - A)^{-1} d\zeta \tag{5.1}$$

would be an inverse for $\zeta I - A$, which is a contradiction. (To see that (5.1) does give an inverse for $\zeta I - A$ recall that $P_i P_j = 0$ if $i \neq j$.)

This example exposes a couple of interesting facts about Banach algebras.

- (1) If $\mathcal{A}' \subset \mathcal{A}$ and $\mathcal{A}'$ has an identity I' , it may happen that I' is not an identity for $\mathcal{A}$
- (2) The spectrum of $A \in \mathcal{A}'$ depends on whether we consider it to be an element of $\mathcal{A}'$ or $\mathcal{A}$.

In the case of the Riesz projections considered last time we have

PROPOSITION 5.1. *If $\mathcal{A}$ is a Banach algebra with identity I and if $P \neq I$ is a projection in $\mathcal{A}$ then for any $A \in \mathcal{A}' = P\mathcal{A}P$*

$$\sigma_{\mathcal{A}}(A) = \sigma_{\mathcal{A}'}(A) \cup \{0\}.$$

PROOF. If $(\zeta I - A)$ is invertible in $\mathcal{A}$ then

$$P(\zeta I - A)^{-1}P(\zeta P - A) = P(\zeta I - A)^{-1}(\zeta I - A)P = P = \cdots = (\zeta P - A)P(\zeta I - A)^{-1}P,$$

so $\zeta P - A$ is invertible in $\mathcal{A}'$. Thus $\sigma_{\mathcal{A}'}(A) \subset \sigma_{\mathcal{A}}(A)$. On the other hand if $\zeta \neq 0$ and $\zeta P - A$ is invertible in $\mathcal{A}'$, with inverse $R'(\zeta)$, then

$$R'(\zeta) + \frac{1}{\zeta}(I - P)$$

is an inverse for $\zeta I - A$. (Note that $(I - P)A = 0$ and $R'(\zeta) = R'(\zeta)P = PR'(\zeta)$.) Thus $\sigma_{\mathcal{A}} \subset \sigma_{\mathcal{A}'}(A) \cup \{0\}$. Finally, since

$$(I - P)A = 0$$

A cannot be invertible (in $\mathcal{A}$) and so $0 \in \sigma_{\mathcal{A}}(A)$. $\square$

What if $\mathcal{A}' \subset \mathcal{A}$ but is not of the form $\mathcal{A}' = P\mathcal{A}P$? In this case, the spectrum can change more dramatically. This can happen even if $\mathcal{A}'$ and $\mathcal{A}$ have a common identity. To see what is going let us consider an example:

Let $\mathcal{A}' =$ uniform closure of polynomials in z in $C(T)$ with $T = \{|z| = 1\}$ and let $\mathcal{A} = C(T)$ (= uniform closure of polynomials in z and $\bar{z}$.) Consider $f(z) = z \in \mathcal{A}' \subset \mathcal{A}$.

PROPOSITION 5.2. (1) $\sigma_{C(T)}(z) = T$.

(2) $\sigma_{\mathcal{A}'}(z) = \overline{\mathbb{D}} = \{|\zeta| \leq 1\}$.

PROOF. For (1), note that if $\zeta \notin T$ then $(\zeta - z)^{-1}$ is a continuous function on T . For (2), first note that if $|\zeta| > 1$ then

$$\frac{1}{\zeta - z} = \sum_{n=0}^{\infty} z^n \frac{1}{\zeta^{n+1}}$$

is absolutely convergent, so $(\zeta - z)^{-1} \in \mathcal{A}'$.

To prove that any $|\zeta| \leq 1$ is in the spectrum of z as an element of $\mathcal{A}'$ let us note that if $f \in \mathcal{A}'$ then there is a sequence p_n of polynomials converging uniformly to f . In particular, p_n is Cauchy in the uniform norm on T . By the maximum principle p_n is a Cauchy sequence in $C(\overline{\mathbb{D}})$. Thus $p_n \rightarrow F$ uniformly on $\overline{\mathbb{D}}$, with F an analytic function such that $F(z) = f(z)$ for $|z| = 1$. Now if $\zeta \in \mathbb{C}$ and there is $f \in \mathcal{A}'$ such that $(\zeta - z)f(z) = 1$ on T then $(\zeta - z)F(z) = 1$ on $\overline{\mathbb{D}}$. Since F is continuous, we must have $|\zeta| > 1$. $\square$

REMARK. We have shown that $\mathcal{A}'$ can be identified with the algebra of continuous functions on $\overline{\mathbb{D}}$ analytic in the interior.

This example is typical in that what can happen is that in the smaller algebra some “holes” in the spectrum can be filled in. To make this precise, let us define

DEFINITION 5.1. If $K \subset \mathbb{C}$, let

$$\|f\|_K = \sup\{|f(z)| : z \in K\}.$$

If K is compact, let $\widehat{K}$ denote the polynomially convex hull of K :

$$\widehat{K} := \{z \in \mathbb{C} : |p(z)| \leq \|p\|_K \text{ for every polynomial } p\}.$$

We say that K is polynomially convex if $\widehat{K} = K$.

PROPOSITION 5.3. *If K is compact in $\mathbb{C}$ then $\mathbb{C} \setminus \widehat{K}$ is the unbounded component of $\mathbb{C} \setminus K$. Thus K is polynomially convex if and only if $\mathbb{C} \setminus K$ is connected.*

PROOF. Let $U_0, U_1, \dots$ be the components of $\mathbb{C} \setminus K$, with U_0 the unbounded component. For $n \geq 1$, U_n is a bounded open set with $\partial U_n \subset K$ (why?). By the maximum principle $U_n \subset \widehat{K}$. Since $K \subset \widehat{K}$ we have

$$L := K \cup \bigcup_{n=1}^{\infty} U_n \subset \widehat{K}.$$

On the other hand if $\zeta \in U_0$ then $(\zeta - z)^{-1}$ is analytic in a neighborhood of L . By Runge's Theorem there is a sequence of polynomials p_n converging uniformly to $(\zeta - z)^{-1}$ on L . Let $q_n(z) = (z - \zeta)p_n(z)$. Then $q_n \rightarrow 1$ uniformly on L . However,

$$1 - q_n(\zeta) = 1,$$

so for sufficiently large n

$$|1 - q_n(\zeta)| > \|1 - q_n\|_K.$$

Thus $\zeta \notin \widehat{K}$. □

THEOREM 5.4. *If $\mathcal{A}' \subset \mathcal{A}$ are Banach algebras with a common identity and $A \in \mathcal{A}'$ then*

- (1) $\sigma_{\mathcal{A}}(A) \subset \sigma_{\mathcal{A}'}(A)$
- (2) $\partial\sigma_{\mathcal{A}'}(A) \subset \partial\sigma_{\mathcal{A}}(A)$
- (3) $\widehat{\sigma}_{\mathcal{A}}(A) = \widehat{\sigma}_{\mathcal{A}'}(A)$
- (4) *If U is a bounded component of $\mathbb{C} \setminus \sigma_{\mathcal{A}}(A)$, then either $U \subset \sigma_{\mathcal{A}'}(A)$ or $U \cap \sigma_{\mathcal{A}'}(A) = \emptyset$.*
- (5) *If $\mathcal{A}' = \text{closure in } \mathcal{A} \text{ of polynomials in } A$, then $\sigma_{\mathcal{A}'}(A) = \widehat{\sigma}_{\mathcal{A}}(A)$.*

PROOF. Let I denote the identity in $\mathcal{A}$ and $\mathcal{A}'$. If $(\zeta I - A)$ is invertible in $\mathcal{A}'$, then since the identity is the same in $\mathcal{A}$ and $\mathcal{A}'$, it is also invertible in $\mathcal{A}$. (1) follows.

Suppose $\zeta \in \partial\sigma_{\mathcal{A}'}(A)$. We must show that $\zeta \in \sigma_{\mathcal{A}}(A)$. (It follows from (1) that then $\zeta \in \partial\sigma_{\mathcal{A}}(A)$.) Suppose on the contrary that there is $R \in \mathcal{A}$ such that

$$R(\zeta I - A) = (\zeta I - A)R = I.$$

Since $\zeta \in \partial\sigma_{\mathcal{A}'}(A)$ there is $\zeta_n \rightarrow \zeta$ with $\zeta_n \in \mathbb{C} \setminus \sigma_{\mathcal{A}'}(A)$. Thus $(\zeta_n I - A)^{-1} \in \mathcal{A}'$. It follows that

$$(\zeta_n I - A)^{-1} \rightarrow R \quad \text{in } \mathcal{A},$$

but since $\mathcal{A}'$ is a Banach space we then have $R \in \mathcal{A}'$. This contradicts the fact that $\zeta \in \sigma_{\mathcal{A}'}(A)$. (2) follows.

(3) follows from (1) and (2) and the maximum principle.

To prove (4), let $G_1 = U \cap \sigma_{\mathcal{A}'}(A)$ and $G_2 = U \setminus \sigma_{\mathcal{A}'}(A)$. So $U = G_1 \cup G_2$ and $G_1 \cap G_2 = \emptyset$. Clearly G_2 is open. Furthermore, since $U \cap \sigma_{\mathcal{A}}(A) = \emptyset$ and $\partial\sigma_{\mathcal{A}'}(A) \subset \sigma_{\mathcal{A}}(A)$ we conclude that $G_1 = U \cap \text{int } \sigma_{\mathcal{A}'}(A)$ is also open. Since U is connected either $G_1 = \emptyset$ or $G_2 = \emptyset$.

Finally, let $\mathcal{A}' = \text{closure of polynomials in } A$. By (1) and (3) we have

$$\sigma_{\mathcal{A}} \subset \sigma_{\mathcal{A}'}(A) \subset \widehat{\sigma}_{\mathcal{A}}(A).$$

On the other hand, if $\zeta \notin \sigma_{\mathcal{A}'}(A)$ then $(\zeta I - A)^{-1} \in \mathcal{A}'$. Thus there is a sequence p_n of polynomials such that

$$p_n(A) \rightarrow (\zeta I - A)^{-1}.$$

Let $q_n(z) = (z - \zeta)p_n(z)$. So $\|q_n(A) - 1\| \rightarrow 0$. By the spectral mapping theorem

$$\sigma_{\mathcal{A}}(q_n(A)) = q_n(\sigma_{\mathcal{A}}(A)).$$

Thus for large enough n

$$1 > \|1 - q_n(A)\| \geq \text{sp-rad}(1 - q_n(A)) = \sup\{|1 - q_n(w)| : w \in \sigma_{\mathcal{A}}(A)\} = \|1 - q_n\|_{\sigma_{\mathcal{A}}(A)}.$$

Since $|1 - q_n(\zeta)| = 1$ it follows that $\zeta \notin \widehat{\sigma}_{\mathcal{A}}(A)$. □

LECTURE 6

Commutative Banach Algebras

Reading: Chapter 18 in Lax We will now specialize to spectral theory in an algebra $\mathcal{A}$ with a unit I and such that the multiplication is commutative:

$$AB = BA \quad \text{for all } A, B \in \mathcal{A}.$$

The theory we will develop here is due to I. M. Gelfand. Throughout this lecture all Banach algebras will be commutative and have a unit.

DEFINITION 6.1. A multiplicative functional p on a Banach algebra $\mathcal{A}$ is a homomorphism of $\mathcal{A}$ into $\mathbb{C}$.

So $p : \mathcal{A} \rightarrow \mathbb{C}$ is a linear functional and

$$p(AB) = p(A)p(B).$$

This definition is purely algebraic. In particular, it is not assumed that p is bounded. However we have

THEOREM 6.1. *Every multiplicative functional p on a commutative Banach algebra is a contraction:*

$$|p(A)| \leq \|A\|.$$

PROOF. Since

$$p(A) = p(IA) = p(I)p(A) \quad \forall A \in \mathcal{A}$$

we have either $p(A) = 0$ for all A or $p(I) = 1$. The first case is trivial.

In the second case, if A is invertible then

$$p(A^{-1})p(A) = p(I) = 1,$$

and so

LEMMA 6.2. *If $p \neq 0$ is a multiplicative functional on a Banach algebra and A is invertible then $p(A) \neq 0$.*

Now suppose $|p(A)| > \|A\|$ for some A . Let

$$B = \frac{A}{p(A)}.$$

Then

$$\|B\| < 1.$$

Thus $I - B$ is invertible, and

$$p(I - B) = p(I) - \frac{p(A)}{p(A)} = 0,$$

which is a contradiction. □

There is a lot of interplay between algebraic and analytic notions in the context of Banach algebras.

DEFINITION 6.2. A subset $\mathcal{I}$ of a commutative Banach algebra $\mathcal{A}$ is called an ideal if

- (1) $\mathcal{I}$ is a linear subspace of $\mathcal{A}$
- (2) $A\mathcal{I} \subset \mathcal{I}$ for any $A \in \mathcal{A}$
- (3) $\mathcal{I} \neq 0, \mathcal{A}$.

Again, this is a purely algebraic notion. The following is a standard algebraic fact:

PROPOSITION 6.3. Let $\mathcal{A}$ and $\mathcal{B}$ be commutative algebras with units and $q : \mathcal{A} \rightarrow \mathcal{B}$ a homomorphism. Suppose that q

- (1) is not an isomorphism, and
- (2) is not the zero map.

Then the

$$\ker q = \{A \in \mathcal{A} : q(A) = 0\}$$

is an ideal in $\mathcal{A}$. Conversely any ideal in $\mathcal{A}$ is the kernel of a homomorphism satisfying (1) and (2)

SKETCH OF PROOF. It is easy to see that $\ker q$ is an ideal. Given an ideal $\mathcal{I}$, to construct the homomorphism, we let $\mathcal{B} = \mathcal{A}/\mathcal{I}$. That is

$$\mathcal{B} = \{\text{equivalence classes for } A \sim B \text{ iff } A - B \in \mathcal{I}\}.$$

Check that $\mathcal{B}$ is an algebra with addition or multiplication given by addition or multiplication of any pair of representatives. Now let $q : \mathcal{A} \rightarrow \mathcal{B}$ be the map

$$q(A) = [A] = \text{equivalence class containing } A. \quad \square$$

An ideal can contain no invertible elements. Indeed if A is invertible and A were in $\mathcal{I}$, then $A^{-1}A = I$ would be in $\mathcal{I}$ which would imply $AI = A \in \mathcal{I}$ for all A , that is $\mathcal{I} = \mathcal{A}$. On the other hand

LEMMA 6.4. Every non-invertible element B of $\mathcal{A}$ belongs to an ideal.

PROOF. If $B = 0$ it is in every ideal. (Ideals are, in particular, vector spaces.) If $B \neq 0$ then $B\mathcal{A} = \{BA : A \in \mathcal{A}\}$ is an ideal and contains B .

EXERCISE 13. Show that $B\mathcal{A}$ is an ideal if B is not invertible.

□

DEFINITION 6.3. A maximal ideal is an ideal that is not contained in a larger ideal.

The space of ideals in $\mathcal{A}$ can be partially ordered by inclusion. It is easy to see that the union of an arbitrary collection of ideals is itself an ideal. Thus Zorn's lemma gives

LEMMA 6.5. Every ideal is contained in a maximal ideal. In particular, every non-invertible element $B \in \mathcal{A}$ belongs to a maximal ideal.

LEMMA 6.6. Let $\mathcal{M}$ be a maximal ideal in $\mathcal{A}$. Every non-zero element of $\mathcal{A}/\mathcal{M}$ is invertible.

REMARK. That is, $\mathcal{A}/\mathcal{M}$ is a division algebra.

PROOF. Suppose $[B] \in \mathcal{A}/\mathcal{M}$ is not invertible. Then $[B]\mathcal{A}/\mathcal{M} = (B\mathcal{A})/\mathcal{M}$ is an ideal. Let

$$\mathcal{I} = \{A \in \mathcal{A} : [A] \in [B]\mathcal{A}/\mathcal{M}\},$$

that is

$$\mathcal{I} = \{A \in \mathcal{A} : A = BK + M \text{ for some } K \in \mathcal{A} \text{ and } M \in \mathcal{M}\}.$$

EXERCISE 14. Show that $\mathcal{I}$ is an ideal.

Since $\mathcal{I}$ is an ideal and clearly $\mathcal{I} \supset \mathcal{M}$, we must have $\mathcal{M} = \mathcal{I}$. Since $B = BI + 0 \in \mathcal{I}$, it follows that $B \in \mathcal{M}$. That is, $[B] = 0$. $\square$

So far we have done no analysis on ideals. To proceed we need an analytic result:

THEOREM 6.7 (Mazur). *Let $\mathcal{A}$ be a Banach algebra with unit that is a division algebra. Then $\mathcal{A}$ is isomorphic to $\mathbb{C}$.*

PROOF. Let $B \in \mathcal{A}$. The spectrum of B is non-empty. Thus there is $\zeta \in \mathbb{C}$ such that $\zeta I - B$ is non-invertible. Since $\mathcal{A}$ is a division algebra, $\zeta I = B$. Thus every element of $\mathcal{A}$ is a multiple of the identity. The map $B \rightarrow \zeta$ is the isomorphism onto $\mathbb{C}$. $\square$

We would like to conclude from Mazur's theorem that $\mathcal{A}/\mathcal{M} \cong \mathbb{C}$ for any maximal ideal $\mathcal{M}$. Indeed, we have seen that $\mathcal{A}/\mathcal{M}$ is a division algebra. However, we are not done as we have not shown it is a Banach algebra. (There are division algebras not isomorphic to $\mathbb{C}$. For example, the algebra of rational functions on $\mathbb{C}$.)

To show that $\mathcal{A}/\mathcal{M}$ is a Banach algebra, we must show in particular that it is a Banach space. That this is true follows because

LEMMA 6.8. *Let $\mathcal{I}$ be an ideal in a commutative Banach algebra. Then the closure $\overline{\mathcal{I}}$ of $\mathcal{I}$ is an ideal. In particular, a maximal ideal $\mathcal{M}$ is closed.*

EXERCISE 15. Prove this lemma.

Thus $\mathcal{A}/\mathcal{M}$ is a quotient of Banach spaces. It follows that it is a Banach space in the following norm:

$$\|[B]\| = \inf_{M \in \mathcal{M}} \|B + M\|.$$

(See Chapter 5.)

LEMMA 6.9. *Let $\mathcal{I}$ be a closed ideal in a commutative Banach algebra $\mathcal{A}$. Then $\mathcal{A}/\mathcal{I}$ is a Banach algebra.*

EXERCISE 16. Prove this lemma

Thus, given a maximal ideal $\mathcal{M}$, the quotient $\mathcal{A}/\mathcal{M}$ is a Banach division algebra and, thus, naturally isomorphic to $\mathbb{C}$ by Mazur's theorem. In particular, the quotient map

$$p(B) = [B]$$

is a multiplicative functional. In fact,

THEOREM 6.10. *Let $\mathcal{A}$ be a commutative Banach algebra. There is a one-to-one correspondence between non-zero multiplicative functionals and maximal ideals given by*

$$\mathcal{M} \mapsto p_{\mathcal{M}}(B) = [B], \quad p_{\mathcal{M}} : \mathcal{A} \rightarrow \mathcal{A}/\mathcal{M} \cong \mathbb{C},$$

and

$$p \mapsto \ker p.$$

PROOF. We have already seen that the quotient map associated to any maximal ideal is a multiplicative functional, so it remains to show that $\ker p$ is a maximal ideal for any multiplicative functional. This is a general algebraic fact. Since p is a non-zero linear functional, $\ker p$ is a subspace of co-dimension 1. Thus any subspace $\mathcal{V} \supset \ker p$ satisfies $\mathcal{V} = \mathcal{A}$ or $\mathcal{V} = \ker p$. Since any ideal $\mathcal{M} \supset \ker p$ is a subspace with $\mathcal{M} \neq \mathcal{A}$, we conclude that $\mathcal{M} = \ker p$ is a maximal ideal. $\square$

COROLLARY 6.11. *An element B of a commutative Banach algebra with unit is invertible if and only if*

$$p(B) \neq 0$$

for all multiplicative functionals.

PROOF. We have already seen that B invertible $\implies p(B) \neq 0$ for all m.f.'s.

Conversely, if B is singular it is contained in a maximal ideal $\mathcal{M}$. Then $p_{\mathcal{M}}(B) = 0$. $\square$

LECTURE 7

Spectral theory in commutative Banach Algebras

THEOREM 7.1. *Let $\mathcal{A}$ be a commutative Banach algebra and let $B \in \mathcal{A}$. Then*

$$\sigma(B) = \{p(B) : p \text{ is a multiplicative linear functional}\}.$$

PROOF. $\zeta \in \sigma(B)$ if and only if $\zeta I - B$ is not invertible. Last time, we saw that this happens if and only if

$$p(\zeta I - B) = 0$$

for some multiplicative functional p . That is if and only if

$$\zeta = p(B). \quad \square.$$

The set $\mathcal{J} = \{\text{maximal ideals in } \mathcal{A}\}$ is called the spectrum of the algebra $\mathcal{A}$. Using the correspondence $\mathcal{M} \sim p_{\mathcal{M}}$ between maximal ideals and multiplicative functionals established last time, we have a natural correspondence between $\mathcal{A}$ and an algebra of functions on $\mathcal{J}$, namely

$$A \mapsto f_A(\mathcal{M}) = p_{\mathcal{M}}(A), \quad (\star)$$

where $p_{\mathcal{M}}$ is the multiplicative functional with kernel $\mathcal{M}$. This map is called the Gelfand representation of $\mathcal{A}$.

THEOREM 7.2.

- (1) *The Gelfand representation is a homomorphism of $\mathcal{A}$ into the algebra of bounded functions on $\mathcal{J}$.*
- (2) *$|f_A(\mathcal{M})| \leq \|A\|$ for all $A \in \mathcal{A}$ and $\mathcal{M} \in \mathcal{J}$.*
- (3) *The spectrum of A is the range of f_A .*
- (4) *The identity I is represented by $f_I = 1$.*
- (5) *The functions f_A separate points of $\mathcal{J}$: given $\mathcal{M}$ and $\mathcal{M}'$ distinct there is $A \in \mathcal{A}$ such that*

$$f_A(\mathcal{M}) \neq f_A(\mathcal{M}').$$

PROOF.

EXERCISE 17. Verify (1), (2), (3), and (4).

To see (5) note that given $A \in \mathcal{M} \setminus \mathcal{M}'$ we have

$$f_A(\mathcal{M}) = 0 \quad \text{and} \quad f_A(\mathcal{M}') \neq 0.$$

□

DEFINITION 7.1. The natural topology on $\mathcal{J}$ is the weakest topology in which all the functions f_A , $A \in \mathcal{M}$, are continuous. It is called the Gelfand topology.

THEOREM 7.3. *$\mathcal{J}$ is a compact Hausdorff space in the Gelfand topology.*

PROOF. This is a standard proof based on Tychonoff's theorem. Let

$$P = \prod_{A \in \mathcal{A}} D_{\|A\|},$$

with $D_{\|A\|}$ the closed disk of radius $\|A\|$ in $\mathbb{C}$. By Tychonoff's theorem P is compact in the product topology. By part (2) of the first theorem

$$f_A(\mathcal{M}) \in D_{\|A\|},$$

so

$$\Phi(\mathcal{M})_A = f_A(\mathcal{M})$$

defines a map from $\mathcal{J} \rightarrow P$. By (5) this map is injective.

EXERCISE 18. Check that the Gelfand topology is the same as the topology induced on $\mathcal{J}$ by this embedding.

Since P is compact, it suffices to show that $\Phi(\mathcal{J})$ is closed.

EXERCISE 19. Show that $\Phi(\mathcal{J})$ is closed. Namely, show that any point $t = t.$ in the closure of $\Phi(\mathcal{J})$ is a homomorphism

$$t_{A+cB} = t_A + ct_B \quad \text{and} \quad t_{AB} = t_A t_B.$$

The Hausdorff property for $\mathcal{J}$ follows since f_A separate points. □

The Gelfand representation need not be injective. For example

$$\mathcal{A} = \left\{ \begin{pmatrix} z & w \\ 0 & z \end{pmatrix} : z, w \in \mathbb{C} \right\}$$

is a commutative Banach algebra, with identity. (Use the matrix norm, or

$$\left\| \begin{pmatrix} z & w \\ 0 & z \end{pmatrix} \right\| = |z| + |w|,$$

which is also sub-multiplicative.) It has a unique maximal ideal namely

$$\mathcal{M} = \left\{ \begin{pmatrix} 0 & w \\ 0 & 0 \end{pmatrix} : w \in \mathbb{C} \right\}.$$

The Gelfand homomorphism is the map

$$\begin{pmatrix} z & w \\ 0 & z \end{pmatrix} \rightarrow z, \quad \mathcal{A} \rightarrow \mathbb{C}.$$

In general, the kernel of the Gelfand representation is

$$\mathcal{R} = \bigcap_{\mathcal{M} \in \mathcal{J}} \mathcal{M},$$

which is called the radical of $\mathcal{A}$.

PROPOSITION 7.4. $A \in \mathcal{R}$ if and only if $\sigma(A) = \{0\}$.

PROOF. This follows from the identity

$$\sigma(A) = \{p_{\mathcal{M}}(A) : \mathcal{M} \in \mathcal{J}\}. \quad \square$$

In particular, $\mathcal{R}$ contains all the nilpotent elements (if there are any). More generally, if $\|A^n\|^{\frac{1}{n}} \rightarrow 0$, so $\text{sp-rad}(A) = 0$, then $A \in \mathcal{R}$.

PROPOSITION 7.5. $\mathcal{R}$ is closed, and is an ideal if $\mathcal{R} \neq 0$.

EXERCISE 20. Prove this.

The radical is essentially the barrier to representing $\mathcal{A}$ as an algebra of functions. Since $\mathcal{R}$ is closed, we may consider the quotient Banach algebra $\mathcal{A}/\mathcal{R}$, which has trivial radical. The Gelfand representation shows that:

THEOREM 7.6. *If $\mathcal{A}$ is a commutative Banach Algebra, then there is a compact Hausdorff space Ω and continuous (bounded) injective homomorphism of $\mathcal{A}/\mathcal{R}$ into $C(\Omega)$.*

Thus a commutative Banach Algebra with trivial radical may be thought of as a subalgebra of the continuous functions on a compact Hausdorff space, and in fact the algebra determines the space.

LECTURE 8

C^* algebras

The algebra of functions on a compact Hausdorff space has an additional structure — complex conjugation — which is not present in commutative Banach algebras. What happens if we put it there?

More generally, we can define

DEFINITION 8.1. A $*$ -operation on a Banach algebra $\mathcal{A}$ is a map $A \mapsto A^*$ from $\mathcal{A} \rightarrow \mathcal{A}$ satisfying

- (1) $(A^*)^* = A$
- (2) $(AB)^* = B^*A^*$
- (3) $(A + B)^* = A^* + B^*$
- (4) $(wA)^* = \overline{w}A^*$.

A C^* algebra $\mathcal{A}$ is a Banach algebra together with a $*$ -operation such that

$$\|A\|^2 = \|A^*A\|.$$

The prime example of a C^* algebra is the algebra of bounded operators on a Hilbert space. In fact, although we will not show this, any C^* algebra is isometrically isomorphic to a sub algebra of the bounded operators on a Hilbert space. A second example is $C_0(\Omega)$ with Ω a locally compact topological space. This example is commutative. If Ω is compact then $C_0(\Omega) = C(\Omega)$ has an identity. If Ω is locally compact then $C_0(\Omega)$ does not have an identity.

PROPOSITION 8.1. If $\mathcal{A}$ is a C^* algebra, then $\|A\| = \|A^*\|$.

PROOF. Note that $\|A\|^2 = \|A^*A\| \leq \|A^*\| \|A\|$, so $\|A\| \leq \|A^*\|$. □

DEFINITION 8.2. An element of a C^* algebra is self-adjoint if $A^* = A$, is anti-self-adjoint if $A^* = -A$ and is unitary if $A^*A = I$.

THEOREM 8.2. If p is a multiplicative functional on a C^* algebra then

- (1) $p(A) \in \mathbb{R}$ if A is self-adjoint.
- (2) $p(A^*) = \overline{p(A)}$.
- (3) $p(A^*A) \geq 0$.
- (4) $|p(U)| = 1$ if U is unitary.

PROOF. We already have $\|p\| \leq 1$ (the proof given above goes through for general Banach algebras). Let A is self adjoint and

$$p(A) = a + ib.$$

With $T_t = A + itI$, we have

$$a^2 + (b + t)^2 = |p(T_t)|^2 \leq \|T_t\|^2.$$

On the other hand, $T_t^* = A - itI$ so

$$T_t^*T_t = A^2 + t^2I,$$

and

$$\|T_t\|^2 \leq \|A\|^2 + t^2.$$

Thus

$$a^2 + (b + t)^2 \leq \|A\|^2 + t^2.$$

This inequality can hold for all t if and only if $b = 0$. So $p(A) = a \in \mathbb{R}$, and (1) follows.

For general A we may write

$$A = \frac{1}{2}(A + A^*) + i\frac{1}{2}(A - A^*),$$

and

$$A^* = \frac{1}{2}(A + A^*) - i\frac{1}{2}(A - A^*),$$

so (2) follows from (1). (3) and (4) follow from (2) since $p(A^*A) = p(A^*)p(A)$. $\square$

COROLLARY 8.3. *If $\mathcal{A}$ is a C^* algebra then*

- (1) *If A is self adjoint $\sigma(A) \subset \mathbb{R}$.*
- (2) *If A is anti-self-adjoint $\sigma(A) \subset i\mathbb{R}$*
- (3) *If A is unitary $\sigma(A) \subset \{|z| = 1\}$.*

PROOF. Use the Gelfand theory

$$\sigma(A) = \{p(A) : p \text{ is a multiplicative functional}\}.$$

$\square$

THEOREM 8.4. *Let $\mathcal{A} \subset \mathcal{B}$ be C^* algebras with the same identity and norm. If $A \in \mathcal{A}$ then $\sigma_{\mathcal{A}}(A) = \sigma_{\mathcal{B}}(A)$.*

PROOF. First let A be self-adjoint and let $\mathcal{C}$ = the algebra generated by A . So $\mathcal{C}$ is a commutative C^* algebra and $\mathcal{C} \subset \mathcal{A} \subset \mathcal{B}$. Since $\mathcal{C}$ is commutative, we have $\sigma_{\mathcal{C}}(A) \subset \mathbb{R}$. Thus

$$\sigma_{\mathcal{B}}(A) \subset \sigma_{\mathcal{A}}(A) \subset \sigma_{\mathcal{C}}(A) = \partial\sigma_{\mathcal{C}}(A) \subset \partial\sigma_{\mathcal{A}}(A) \subset \partial\sigma_{\mathcal{B}}(A).$$

It follows that $\sigma_{\mathcal{B}}(A) = \sigma_{\mathcal{A}}(A) = \sigma_{\mathcal{C}}(A)$.

To prove the general statement, it suffices to show that if A is invertible in $\mathcal{B}$ it is invertible in $\mathcal{A}$. So suppose we have $B \in \mathcal{B}$ such that $BA = AB = I$. It follows that

$$(A^*A)(BB^*) = (BB^*)(A^*A) = I.$$

Since A^*A is self-adjoint, the first part of the proof implies that A^*A is invertible in $\mathcal{A}$. Thus $BB^* \in \mathcal{A}$. Thus

$$B = B(B^*A^*) = (BB^*)A^* \in \mathcal{A}. \quad \square$$

Thus in any C^* algebra we can use the Gelfand theory to compute the spectrum, because the spectrum of an element A is the same as its spectrum in the smallest C^* algebra containing it

$$C^*(A) = \text{algebra generated by } A \text{ and } A^*.$$

DEFINITION 8.3. An element A in a C^* algebra is called normal if $AA^* = A^*A$.

THEOREM 8.5. *For normal A in a C^* algebra,*

$$\text{sp-rad}(A) = \|A\|. \quad (*)$$

In particular, in a commutative C^ algebra $(*)$ holds for every A .*

PROOF. For self-adjoint A we have

$$\|A^2\| = \|A\|^2.$$

It follows that

$$\|A^{2^k}\| = \|A\|^{2^k}.$$

Thus

$$\text{sp-rad}(A) = \lim_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}} = \lim_{k \rightarrow \infty} \|A^{2^k}\|^{2^{-k}} = \|A\|.$$

If A is normal, then $C^*(A, A^*)$ is commutative, so by the Gelfand theory we have

$$\|A\|^2 = \|A^*A\| = \text{sp-rad}(A^*A) = \sup_p p(A^*A) = \sup_p |p(A)|^2 = \text{sp-rad}(A)^2,$$

where the sup is over multiplicative functionals on $C^*(A, A^*)$. $\square$

COROLLARY 8.6. *If $\mathcal{A}$ is a commutative C^* algebra with unit then there is a compact Hausdorff space Ω and an isometric isomorphism $\Phi : \mathcal{A} \rightarrow C(\Omega)$.*

PROOF. Let Ω be the maximal ideal space of $\mathcal{A}$ in the Gelfand topology, and let $\Phi : \mathcal{A} \rightarrow C(\Omega)$ be the Gelfand representation. By the previous theorem, $\text{sp-rad}(A) = 0 \implies A = 0$ so the radical of $\mathcal{A}$ is $\{0\}$. Thus the Gelfand representation is injective. Also,

$$\|A\|^2 = \sup_{\mathcal{M}} |p_{\mathcal{M}}(A)|^2 = \sup_{\mathcal{M}} |\Phi(A)(\mathcal{M})|^2,$$

so the Gelfand representation is an isometry.

It remains to show that the range of Φ is all of $C(\Omega)$. This follows from Stone-Weierstrass since $\Phi(\mathcal{A})$ is closed, separates points, and is closed under conjugation. $\square$

Part 2

Spectral Theory for operators on a Hilbert space

LECTURE 9

Self-adjoint operators

Reading: Chapter 31

Recall that an operator $A \in \mathcal{L}(H)$, H a complex Hilbert space, is self-adjoint (or Hermitian or symmetric) if

$$\langle Ax, y \rangle = \langle x, Ay \rangle \quad \text{for all } x, y \in H.$$

We will assume unless mentioned otherwise that H is separable. So H is isomorphic either to $\mathbb{C}^n$ or ℓ^2 .

Last term we saw that if A is self-adjoint and compact then there is a orthonormal basis of eigenvectors for A :

$$Ae_n = \lambda_n e_n, \quad \langle e_n, e_m \rangle = \delta_{n,m}, \quad \overline{\text{span}\{e_n : n = 1, \dots\}} = H.$$

Thus given $x \in H$ we have

$$x = \sum_n a_n e_n, \quad Ax = \sum_n a_n \lambda_n e_n,$$

with $a_n = \langle x_n, e_n \rangle$.

An expansion into eigenvectors does not exist for an arbitrary self-adjoint operator. For instance multiplication by x in $L^2(-1, 1)$,

$$Mf(x) = xf(x)$$

has no eigenvectors in L^2 . In this case there are eigenvectors in the sense of distributions:

$$M\delta(x - \lambda) = \lambda, \quad \lambda \in (-1, 1).$$

In some sense we have

$$f(x) = \int_{-1}^1 f(\lambda) \delta(x - \lambda) d\lambda, \quad Mf(x) = \int_{-1}^1 f(\lambda) \lambda \delta(x - \lambda) d\lambda,$$

analogous to the above expression.

Thus for a general self-adjoint operator the eigenvectors may be quite singular. A slightly better object to work with are projections onto the spaces spanned by eigenvectors with eigenvalues in some set S . In the case of $Mf(x) = xf(x)$, we have

$$E_M(S)f(x) = \chi_S(x)f(x),$$

which is a non-zero projection if S is a set of positive Lebesgue measure in $L^2(-1, 1)$.

Let us rewrite, the expressions for the compact operator A in this form. Let $E(\{\lambda_n\})$ for the projection onto the (finite dimensional) subspace of eigenvectors with eigenvalue λ_n . For a general set $S \subset \mathbb{R}$, let

$$E(S) = \sum_{\lambda_n \in S} E(\{\lambda_n\}).$$

PROPOSITION 9.1.

- (1) $E(S)$ is a projection for each $S \subset \mathbb{R}$.
 (2) $E(S)E(S') = E(S \cap S')$. In particular, if $S \cap S' = \emptyset$ then $\text{ran } E(S) \perp \text{ran } E(S')$.
 (3) If $S_1, \dots, S_n$ are disjoint then

$$E(S_1 \cup \dots \cup S_n) = E(S_1) + \dots + E(S_n).$$

- (4) If $S_j, j = 1, \dots, \infty$, are disjoint then

$$E(\cup_j S_j)x = \sum_{j=1}^{\infty} E(S_j)x$$

for each $x \in H$.

- (5) The same properties hold for $E_M(S)$ provided we restrict our attention to Lebesgue measurable sets.

EXERCISE 21. Prove this.

The maps $S \mapsto E(S)$, $E_M(S)$ are projection valued measures. Note that

$$A = \int \lambda dE((-\infty, \lambda]) \quad M = \int \lambda dE_M((-\infty, \lambda]),$$

with the integrals understood as Stieltje's integrals in the strong operator topology. That is, for every $x \in H$,

$$Ax = \lim_{n \rightarrow \infty} \sum_{j=1}^n \lambda_j^{(n)} E(-\lambda_{j-1}^{(n)}, \lambda_j^{(n)})x,$$

with $\lambda_j^{(n)}$ a partition of the interval $[-\|A\|, \|A\|]$, say, with mesh size $\rightarrow 0$ as $n \rightarrow \infty$.

DEFINITION 9.1. A projection valued measure over H is a map $E : \Sigma \rightarrow \mathcal{L}(H)$ defined on a sigma algebra of sets on some measurable space with the following properties.

- (1) $E(S)$ is an orthogonal projection for every S .
 (2) (finite additivity) $E(S_1) + E(S_2) + \dots + E(S_n) = E(S_1 \cup \dots \cup S_n)$ if S_j are disjoint.
 (3) (strong countable additivity) If $S_j, j = 1, \dots, \infty$, are disjoint then

$$E(\cup_j S_j)x = \sum_{j=1}^{\infty} E(S_j)x$$

for each $x \in H$.

EXERCISE 22. Derive $E(S)E(S') = E(S \cap S')$ from (1) and (2).

Associated to any projection valued measure on $\mathbb{R}$, $E(S)$, with compact support ($E(\mathbb{R} \setminus [-r, r]) = 0$ for some r) there is a bounded self-adjoint operator

$$A = \int \lambda dE((-\infty, \lambda]).$$

Our ultimate goal is to show that the converse is true. This is the “spectral theorem.” (If E does not have compact support, there is still a self-adjoint operator, but it is unbounded. We will get to this.)

That is where we are headed, but it will take a little while.

THEOREM 9.2. *The spectrum of a bounded, self-adjoint operator M on a Hilbert space is a compact subset of the real line and $\text{sp-rad}(M) = \|M\|$.*

PROOF. This follows from the results on C^* algebras. $\square$

THEOREM 9.3. *The spectrum $\sigma(M)$ of a bounded, self adjoint operator M lies in the closed interval $[a, b]$, where*

$$a = \inf_{\|x\|=1} \langle x, Mx \rangle, \quad \text{and} \quad b = \sup_{\|x\|=1} \langle x, Mx \rangle,$$

and $a, b \in \sigma(M)$.

PROOF. If $\lambda < a$ then

$$(b - \lambda) \|x\|^2 \geq \langle x, (M - \lambda I)x \rangle \geq (a - \lambda) \|x\|^2.$$

It follows that $\langle x, (M - \lambda I)y \rangle = \langle x, y \rangle_\lambda$ is an inner product on H , which gives rise to a norm equivalent to $\|\cdot\|$. By the Riesz theorem, for any $y \in H$ the linear functional $\ell(x) = \langle x, y \rangle$ can be represented

$$\langle x, y \rangle = \langle x, z \rangle_\lambda = \langle x, (M - \lambda I)z \rangle$$

for some z . Clearly, the map $y \mapsto z$ is inverse to $z \mapsto (M - \lambda I)z$. Since

$$\|z\|^2 = \langle z, z \rangle \leq \frac{1}{a - \lambda} \langle z, z \rangle_\lambda = \frac{1}{a - \lambda} \langle z, y \rangle \leq \|z\| \|y\|$$

we find that

$$\|z\| \leq \frac{1}{a - \lambda} \|y\|,$$

so the inverse is a bounded map. Thus λ is in the resolvent set of M . A similar argument, with some minus signs, works for b . Thus $\sigma(M) \subset [a, b]$.

Since $\sigma(M) \subset [a, b]$, and $|\langle x, Mx \rangle| \leq \|M\| \|x\|$,

$$\text{sp-rad}(M) \leq \max |a|, |b| \leq \|M\|.$$

Since $\text{sp-rad}(M) = \|M\|$ we see that the larger of a and b in magnitude lies in the spectrum of M . Applying this to $M + cI$ with $c = -a$ and $c = -b$ we conclude that

$$b - a \in \sigma(M - aI) = \sigma(M) - a \implies b \in \sigma(M),$$

and

$$a - b \in \sigma(M - bI) \implies a \in \sigma(M). \quad \square$$

LECTURE 10

Functional calculus and polar decomposition

The Gelfand representation gives an isometric isomorphism from $C^*(A) \rightarrow C(\sigma(A))$, with A a self-adjoint operator. Let us see this in a direct way.

If q is a polynomial, $q(\lambda) = a_n \lambda^n + \cdots a_0$, then we have

$$q(A) = a_n A^n + \cdots a_0 I.$$

By the spectral mapping theorem

$$\sigma(q(A)) = q(\sigma(A)).$$

If the coefficients of q are real then $q(A)$ is self-adjoint and we conclude from Theorem 2 of the last lecture that

$$\|q(A)\| = \sup_{\lambda \in \sigma(A)} |q(\lambda)|. \quad (\star)$$

By the Weierstrass theorem, for any $f \in C(\sigma(A))$ is real valued we can find a sequence of polynomials q_n with real coefficients such that

$$\sup_{\lambda \in \sigma(A)} |f(\lambda) - q_n(\lambda)| \rightarrow 0.$$

It follows from $(\star)$ that $q_n(A)$ is a Cauchy sequence in $\mathcal{L}(H)$. Thus $q_n(A)$ has a limit. We call this limit $f(A)$.

EXERCISE 23. Show that the value of $f(A)$ does not depend on the approximating sequence of polynomials, and that $f(A)$ is self-adjoint. (Recall that f is real valued.)

The map $f \mapsto f(A)$ is called the continuous functional calculus for the self-adjoint operator A . So far we have defined it only for real valued functions, however it extends to complex valued functions applying the real functional calculus to the real and imaginary parts. Thus we have

THEOREM 10.1. *Let $A \in \mathcal{L}(H)$ be self-adjoint. There is a unique isometric algebra homomorphism $A \mapsto f(A)$ from $C(\sigma(A)) \rightarrow \mathcal{L}(H)$ such that $1 \mapsto I$ and $x \mapsto A$. This map satisfies*

- (1) $\|f(A)\| = \sup_{\lambda \in \sigma(A)} |f(\lambda)|$
- (2) $\sigma(f(A)) = f(\sigma(A))$
- (3) $f(A)$ is normal for all $f \in C(\sigma(A))$ and self-adjoint if and only if f is real valued.

One consequence of the functional calculus is the polar decomposition for arbitrary operators. Recall that the polar decomposition of a matrix M is the factorization

$$M = T|M|,$$

where T is a partial isometry and $|M|$ is a non-negative definite matrix.

DEFINITION 10.1. We say that an operator $A \in \mathcal{L}(H)$ is non-negative if

$$\langle Au, u \rangle \geq 0 \quad \text{for all } u \in H.$$

THEOREM 10.2. An operator A is non-negative iff A is self-adjoint and $\sigma(A) \subset [0, \infty)$.

PROOF. If A is non-negative then, in particular $\langle Au, u \rangle$ is real valued for each u and so $\langle Au, u \rangle = \overline{\langle Au, u \rangle} = \langle u, Au \rangle$. It follows that

$$\begin{aligned} 4 \langle Av, u \rangle &= \langle A(v+u), v+u \rangle - \langle A(v-u), v-u \rangle + i[\langle A(v+iu), (v+iu) \rangle - \langle A(v-iu), (v-iu) \rangle] \\ &= \langle v+u, A(v+u) \rangle - \langle v-u, A(v-u) \rangle + i[\langle v+iu, A(v+iu) \rangle - \langle v-iu, A(v-iu) \rangle] \\ &= 4 \langle v, Au \rangle, \end{aligned}$$

so A is self-adjoint. We saw last time that if A self-adjoint implies

$$\inf \sigma(A) = \inf_{\|x\|=1} \langle Ax, x \rangle.$$

Thus $\inf \sigma(A) \geq 0$.

On the other hand if A is self-adjoint and $\sigma(A) \subset [0, \infty)$ then the function $f(\lambda) = \sqrt{\lambda}$ is in $C(\sigma(A))$. Hence $\sqrt{A}$, defined by the functional calculus, is self-adjoint and satisfies $(\sqrt{A})^2 = A$. Thus

$$\langle Au, u \rangle = \langle \sqrt{A}\sqrt{A}u, u \rangle = \langle \sqrt{A}u, \sqrt{A}u \rangle = \|\sqrt{A}u\|^2 \geq 0. \quad \square$$

In the process we have shown

COROLLARY 10.3. Every positive self-adjoint operator has a positive self-adjoint square root.

and

THEOREM 10.4. An operator $A \in \mathcal{L}(H)$ is self-adjoint if and only if $\langle Au, u \rangle \in \mathbb{R}$ for all $u \in H$.

As a consequence of the existence of square roots, we have

THEOREM 10.5 (Polar decomposition). Every operator $A \in \mathcal{L}(H)$ can be factored as $A = T|A|$ where T is a partial isometry and $|A| = \sqrt{A^\dagger A}$ is non-negative. The polar decomposition has the following properties

- (1) $\ker |A| = \ker A$
- (2) T is an isometry from $\text{ran } |A|$ onto $\text{ran } A$.
- (3) T is not unique, but can be taken to be 0 on $(\text{ran } |A|)^\perp = \ker A$.

PROOF. $A^\dagger A$ is non-negative since

$$\langle A^\dagger Au, u \rangle = \langle Au, Au \rangle \geq 0.$$

Thus $|A|$ is well defined. Since

$$\|Au\|^2 = \langle Au, Au \rangle = \langle A^\dagger Au, u \rangle = \langle |A|^2 u, u \rangle = \langle |A|u, |A|u \rangle = \||A|u\|^2,$$

we see that $Au = 0$ if and only if $|A|u = 0$. Thus $\ker |A| = \ker A$. Furthermore, the map $T : \text{ran } |A| \rightarrow \text{ran } A$ defined by

$$T|A|u = Au$$

is well defined, since if $|A|u = |A|v$ then $u - v \in \ker A$ so $Au = Av$. Clearly this map is an isometry. Extending T to be 0 on $(\operatorname{ran} |A|)^\perp$, the map is a partial isometry. Finally

$$(\operatorname{ran} |A|)^\perp = \ker |A| = \ker A,$$

since for arbitrary B we have $(\operatorname{ran} B)^\perp = \ker B^\dagger$. □

LECTURE 11

Spectral resolution

Given a self-adjoint operator A on a Hilbert space, the functional calculus $f \mapsto f(A)$ is a bounded linear map from $C(\sigma(A)) \rightarrow \mathcal{L}(H)$. As such, the maps

$$\ell_{x,y}(f) = \langle f(A)x, y \rangle,$$

defined for every pair $x, y \in H$, are bounded linear functionals. According to the Riesz representation theorem, then, to each pair $x, y \in H$ there corresponds a complex regular Borel measure on $\sigma(A)$ such that

$$\langle f(A)x, y \rangle = \int_{\sigma(A)} f(\lambda) dm_{x,y}(\lambda).$$

THEOREM 11.1.

- (1) $m_{x,y}$ is sesquilinear in x, y (linear in x and conjugate linear in y).
- (2) $m_{y,x} = \overline{m_{x,y}}$.
- (3) $\|m_{y,x}\| \leq \|x\| \|y\|$.
- (4) The measures $m_{x,x}$ are non-negative.

REMARK. $\|m_{x,y}\|$ denotes the total variation norm of $m_{x,y}$:

$$\|m_{x,y}\| = \sup_{f \in C(\sigma(A)) : |f(\lambda)| \leq 1} \left| \int_{\sigma(A)} f(\lambda) dm_{x,y}(\lambda) \right|.$$

PROOF. (1), (2), and (3) are left as exercises. To prove (4) note that if $f(\lambda) \geq 0$ on $\sigma(A)$, then $\sqrt{f} \in C(\sigma(A))$, so

$$\langle f(A)x, x \rangle = \langle \sqrt{f}(A)\sqrt{f}(A)x, x \rangle = \langle \sqrt{f}(A)x, \sqrt{f}(A)x \rangle \geq 0.$$

Thus

$$\int_{\sigma(A)} f(\lambda) dm_{x,x}(\lambda) \geq 0$$

if $f \geq 0$. It follows from the Riesz theorem that $m_{x,x}$ is a non-negative measure. □

Thus for every Borel set $S \subset \sigma(M)$ we have a quadratic form

$$Q_S(x, y) = m_{x,y}(S),$$

which is bounded ($|Q_S(x, y)| \leq \|x\| \|y\|$), sesquilinear and skew-symmetric under interchange of x and y ($Q_S(x, y) = |Q_S(y, x)|$).

THEOREM 11.2. Associated to any function $B(x, y)$ on $H \times H$ which is bounded, sesquilinear and skew-symmetric, there is a bounded self-adjoint operator M such that

$$B(x, y) = \langle Mx, y \rangle.$$

PROOF. Fix y and consider $B(x, y)$ as a function of x . This is a bounded linear functional on H , so by the Riesz-Frechet theorem there is $w \in H$ such that

$$\langle x, w \rangle = B(x, y).$$

Let $My = w$. Since $\|w\| \leq c\|y\|$, so M is bounded. Self-adjointness of M follows from the skew-symmetry:

$$\langle x, My \rangle = B(x, y) = \overline{B(y, x)} = \overline{\langle y, Mx \rangle} = \langle Mx, y \rangle. \quad \square$$

Thus to each Borel subset $S \subset \sigma(A)$ is associated a bounded self-adjoint operator $E(S)$ so that

$$m_{x,y}(S) = \langle E(S)x, y \rangle.$$

THEOREM 11.3.

- (1) $E(\emptyset) = 0$ and $E(\sigma(A)) = I$.
- (2) If $S \cap T = \emptyset$ then $E(S \cup T) = E(S) + E(T)$.
- (3) $E(S \cap T) = E(S)E(T)$
- (4) Each $E(S)$ is an orthogonal projection, and $\text{ran } E(S) \perp \text{ran } E(T)$ if S, T are disjoint.
- (5) $[E(S), E(T)] = 0$ and $[E(S), A] = 0$ for all S, T .
- (6) $E(S)$ is countably additive in the strong operator topology.

REMARK. So $E(S)$ is a projection valued measure as defined in Lecture 9.

PROOF. Clearly $m_{x,y}(\emptyset) = 0$ for all x, y , so $E(\emptyset) = 0$. Likewise

$$m_{x,y}(\sigma(A)) = \int_{\sigma(A)} dm_{x,y}(\lambda) = \langle Ix, y \rangle = \delta_{x,y},$$

so $E(\sigma(A)) = I$. Part (2) follows from the additivity of the measures $m_{x,y}$.

Note that (3) is equivalent to

$$m_{x,y}(S \cap T) = m_{E(T)x,y}(S) \quad \forall x, y \in H, \quad S, T \subset \sigma(A).$$

This in turn is equivalent to

$$\int_T f(\lambda) dm_{x,y}(\lambda) = \langle f(A)E(T)x, y \rangle \quad \forall x, y \in H, \quad f \in C(\sigma(A)), \quad T \subset \sigma(A),$$

which is equivalent to

$$\int_{\sigma(A)} g(\lambda)f(\lambda) dm_{x,y}(\lambda) = \langle f(A)g(A)x, y \rangle \quad \forall x, y \in H, \quad f, g \in C(\sigma(A)),$$

which holds. Since $E(S)$ is self-adjoint and $E(S)^2 = E(S)$, we see that $E(S)$ is an orthogonal projection. Since $E(S)E(T) = 0$ if $S \cap T = \emptyset$, we conclude that $\text{ran } E(S) \perp \text{ran } E(T)$. The first part of (4) follows from (3).

EXERCISE 24. Show that $[E(S), A] = 0$.

EXERCISE 25. Show that $E(S)$ is countably additive in the strong operator topology.

□

Thus,

THEOREM 11.4. *To each self-adjoint operator A , there corresponds a unique projection valued measure E on $\sigma(A)$ such that*

$$f(A) = \int_{\sigma(A)} f(\lambda) dE,$$

for all $f \in C(\sigma(A))$, with the integral on the r.h.s. a norm convergent Riemann-Stieltjes integral.

PROOF. We have already constructed the P.V.M. The uniqueness follows from the uniqueness in the Riesz theorem. If $I_1, \dots, I_n$ are subsets of $\sigma(A)$ with $\cup_j I_j = \sigma(A)$ and I_j pairwise disjoint then

$$\left\| \sum_j a_j E(I_j) \right\| \leq \max_j |a_j|$$

orthogonality of the ranges of $E(I_j)$. It follows that, for continuous f , the oscillation of f on the partition $I_1, \dots, I_n$

$$\text{Osc}(f, \{I_j\}) = \left\| \sum_j \sup_{x, y \in I_j} |f(x) - f(y)| E(I_j) \right\|$$

converges to zero as the mesh

$$\Delta = \max_j \text{diam}(I_j) \rightarrow 0.$$

In the standard way, one concludes the existence of the integral. □

In particular, we have

$$I = \int_{\sigma(A)} dE, \quad A = \int_{\sigma(A)} \lambda dE.$$

One can refine the spectral resolution a bit further. Every Borel measure on μ on the line $\mathbb{R}$ can be written as a sum of three part:

$$\mu = \mu^{(p)} + \mu^{(sc)} + \mu^{(ac)},$$

where

- (1) $\mu^{(p)}$ is the point measure: $\mu^{(p)}(S) = \sum_{x \in S} \mu(\{x\})$.
- (2) $\mu^{(ac)}$ is the absolutely continuous measure: $d\mu^{(ac)}(\lambda) = \frac{d\mu}{d\lambda} d\lambda$.
- (3) $\mu^{(sc)}$ is everything else, and is singular continuous — it has no atoms (so $F(\lambda) = \mu^{(sc)}(-\infty, \lambda)$ is a continuous function) but is supported on a set of measure 0 (so $F'(\lambda) = 0$ almost everywhere).

Applying this decomposition to the measures $m_{x,y}$ we obtain three distinct projection valued measures $E^{(p)}$, $E^{(sc)}$, and $E^{(ac)}$. The measures are orthogonal to one another, that is

$$H = H^{(p)} \oplus H^{(sc)} \oplus H^{(ac)} \quad H^{(\sharp)} = \text{ran } E^{(\sharp)}(\sigma(A)).$$

EXERCISE 26. Show that $H^{(p)}$ = closed linear span of the eigenvectors of A and that $E^{(p)}(S) = \sum_{\lambda \in S} E(\{\lambda\})$, where the sum runs over the eigenvectors of A and $E(\{\lambda\})$ is the projection onto the corresponding eigenspace.

LECTURE 12

Representation of a self-adjoint operator as a multiplication operator

Let $A \in \mathcal{L}(H)$ be self-adjoint and let $x \in H$. Consider the subspace

$$H_x^\circ = \{f(H)x : f \in C(\sigma(A))\}.$$

In general H_x° is not closed. Let H_x denote its closure. So H_x is everything that can be approximated in terms of polynomials in A times x . It is referred to as the cyclic subspace for A containing x .

PROPOSITION 12.1. H_x is an invariant subspace for A .

PROOF. Let $y \in H_x$. So there are $f_n \in C(\sigma(A))$ such that $f_n(A)y \rightarrow x$. Since A is continuous $Af_n(A)y \rightarrow Ax$. But $Af_n(A) = g_n(A)$ with $g_n(\lambda) = \lambda f_n(\lambda)$. $\square$

THEOREM 12.2. The map $T_x f = f(A)x$ extends from $C(\sigma(A))$ to an isometry from $L^2(\sigma(A); m_{x,x}) \rightarrow H_x$, and

$$AT_x = T_x M_\lambda,$$

where M is multiplication by the independent variable on $L^2(m_{x,x})$:

$$Mf(\lambda) = \lambda f(\lambda).$$

That is, if we restrict A to a cyclic subspace it can be represented as multiplication by λ on $L^2(dm_{x,x}(\lambda))$.

PROOF. $m_{x,x}$ is a Borel measure on $\sigma(A)$. It follows that $C(\sigma(A))$ is dense in $L^2(m_{x,x})$. Consider the L^2 norm of a continuous function f :

$$\int_{\sigma(A)} |f(\lambda)|^2 dm_{x,x}(\lambda) = \langle f(A)^\dagger f(A)x, x \rangle = \langle f(A)x, f(A)x \rangle = \|f(A)x\|^2.$$

So the map T_x is an isometry on a dense subset of L^2 . It follows that T_x can be extended uniquely to all of L^2 as an isometry.

Now let $f \in C(\sigma(A))$. Then $Mf \in C(\sigma(A))$. Thus

$$T_x Mf = (Mf)(A)x = Af(A)x.$$

Since this holds on a dense subset, it extends to L^2 . $\square$

What does this say for matrices? For a self-adjoint matrix, the spectrum is a finite set of points in $\mathbb{R}$. Thus $C(\sigma(A)) \sim \mathbb{C}^n$, with n the number of points in the spectrum. The spectral measure $m_{x,x}$ is an assignment of a non-negative weight to each point of the spectrum. The map T_x can be understood as a change of basis on H_x which diagonalizes A . If some eigenvalues are degenerate then we see by simple dimension counting that H_x is not all of H . Likewise H_x could be a proper subspace just because we choose x in a non-generic way. (What happens if x is an eigenvector?) Thus to obtain the full diagonalization of A we need to repeat the procedure. The same is true for self-adjoint operators.

THEOREM 12.3 (Spectral Theorem). *Let $A \in \mathcal{L}(H)$ be self-adjoint. Then there is a finite or countable family of mutually orthogonal closed subspaces $\{H_j : j = 1, \dots, N\}$ and a finite or countable family of Borel probability measures $\{\mu_j : j = 1, \dots, N\}$ ($N < \infty$ or $N = \infty$) such that*

- (1) H_j is invariant for A for $j = 1, \dots, N$,
- (2) There is an isometry $T_j : L^2(\mu_j) \rightarrow H_j$ such that

$$AT_j = T_j M_j,$$

with $M_j f(\lambda) = \lambda f(\lambda)$.

COROLLARY 12.4. *Let $A \in \mathcal{L}(H)$ be self-adjoint. Then there is a measure space Ω, μ and an isometry $T : L^2(\mu) \rightarrow H$ such that*

$$AT = TM_\Phi,$$

with

$$M_\Phi f(\omega) = \Phi(\omega) f(\omega),$$

where Φ is a real valued function in $L^\infty(\mu)$.

PROOF OF COROLLARY. Let μ_j be as in the Theorem. Clearly μ_j are all supported in $\sigma(A)$. Let $\Omega = \sigma(A) \times \mathbb{N}$ with the measure

$$\mu(S) = \sum_j \mu_j(\pi_j(S)),$$

where

$$\pi_j(S) = \{\lambda : (\lambda, j) \in S\}.$$

Now let

$$Tf = \sum_j T_j \Pi_j f,$$

where

$$\Pi_j f(\lambda) = f(\lambda, j).$$

Note that $\Pi_j : L^2(\mu) \rightarrow L^2(\mu_j)$ is a bounded linear map, and

$$\|f\|_{L^2(\mu)}^2 = \sum_j \|\Pi_j f\|_{L^2(\mu_j)}^2.$$

Since $\text{ran } T_j \perp \text{ran } T_k$, it follows that T is an isometry. Also

$$ATf = \sum_j AT_j \Pi_j f = \sum_j T_j M_j \Pi_j f,$$

which shows that $AT = TM_\Phi$ with

$$\Phi(\lambda, j) = \lambda. \quad \square$$

PROOF OF THEOREM. Zornify....

□

LECTURE 13

Spectrum, Spectral Measures and Spectral Multiplicities

Reading: §31.5, but there are some minor errors in Lax's approach. The approach taken in these notes is more general.

Let us take a closer look at the spectral representation obtained last time:

THEOREM (Spectral Theorem). *Let $A \in \mathcal{L}(H)$ be self-adjoint. Then there is a finite or countable family of mutually orthogonal closed subspaces $\{H_j : j = 1, \dots, N\}$ and a finite or countable family of Borel probability measures $\{\mu_j : j = 1, \dots, N\}$ ($N < \infty$ or $N = \infty$) such that*

- (1) H_j is invariant for A for $j = 1, \dots, N$,
- (2) There is an isometry $T_j : L^2(\mu_j) \rightarrow H_j$ such that

$$AT_j = T_j M_j,$$

with $M_j f(\lambda) = \lambda f(\lambda)$.

DEFINITION 13.1. We will call a triple $\{H_j, \mu_j, T_j\}$ as in the theorem a spectral representation of A .

The first thing to note is that there are many choices for the measures. For instance if $\mu_j(S) > 0$ and $\mu_j(T) > 0$ with $S \cap T = \emptyset$ we could split μ_j in two pieces both of which appear in the decomposition. Likewise it could happen that two measures have disjoint support, in which case we could add them together.

To get an idea what happens, think about a compact operator (or a matrix)? Then we have a list of eigenvalues λ_α and multiplicities m_α . The possible spectral measures μ_j we could obtain are just weighted sums of eigenvalues:

$$\mu(\lambda) = \sum_{\alpha} w_{\alpha} \delta(\lambda - \lambda_{\alpha}).$$

An eigenvalue must appear in m_α spectral measures in the list. For instance, if $m_\alpha = 1$ for all α then we can take $N = 1$ with the single spectral measure

$$\mu(\lambda) = \sum_{\alpha} 2^{-\alpha} \delta(\lambda - \lambda_{\alpha}).$$

However, we could take $N = \infty$ with

$$\mu_j(\lambda) = \delta(\lambda - \lambda_j),$$

so $L^2(\mu_j) \cong \mathbb{C}$. What remains the same in all spectral representations are the eigenvalues and the multiplicities: the eigenvalues are the support of the spectral measures and the multiplicity of an eigenvalue is the number of spectral measures in which it appears.

DEFINITION 13.2. The closed support F of a non-negative measure μ is the smallest closed set such that $\mu(F^c) = 0$.

The trouble with the support of a measure is that it gives us very little information about the nature of the measure.

EXAMPLE. Let $\mu(\lambda) = \sum_j 2^{-j} \delta(\lambda - q_j)$, with q_j an enumeration of the rationals in $[0, 1]$. Then the closed support of μ is $[0, 1]$.

Nonetheless,

THEOREM 13.1. *The spectrum of a self-adjoint operator A is*

$$\sigma(A) = \overline{\cup_j \text{closed support of } \mu_j}.$$

EXERCISE 27. Prove this.

THEOREM 13.2. *Let H_j , μ_j , T_j be a spectral representation of A . Let $x \in H$. Then $m_{x,x} \ll \sum_j 2^{-j} \mu_j$.*

PROOF. First note that $\mu_k \ll \sum_j 2^{-j} \mu_j =: \mu$ for each k . By Radon-Nikodym, there are functions $F_k \in L^2(\mu)$ such that

$$d\mu_k(\lambda) = F_k(\lambda) d\mu(\lambda).$$

Now, $x = \sum_j x_j$ with $x_j \in H_j$. Thus

$$\langle x, f(H)x \rangle = \sum_j \langle x_j, f(H)x_j \rangle = \sum_j \int f(\lambda) |f_j(\lambda)|^2 d\mu_j(\lambda) = \int f(\lambda) \sum_j |f_j(\lambda)|^2 F_j(\lambda) d\mu(\lambda),$$

where $f_j = T_j^\dagger x_j \in L^2(\mu_j)$. Since $F_j \geq 0$ and

$$\int \sum_j |f_j(\lambda)|^2 F_j(\lambda) d\mu(\lambda) = \|x\|^2,$$

$\sum_j |f_j(\lambda)|^2 F_j(\lambda) \in L^1(\mu)$ and the result follows. $\square$

COROLLARY 13.3. *If H_j , μ_j , T_j and H'_j , μ'_j , T'_j are two spectral representations of A then $\sum_j 2^{-j} \mu_j$ and $\sum_j 2^{-j} \mu'_j$ are mutually absolutely continuous.*

PROOF. Note that $\sum_j 2^{-j} \mu_j$ is the spectral measure associated to $\sum_j 2^{-\frac{j}{2}} T_j 1$. $\square$

DEFINITION 13.3. The null-class $\mathcal{N}(\mu)$ of a measure μ is the collection of sets S such that $S \subset S'$ with S' measurable and $\mu(S') = 0$. Abstractly, a collection of sets $\mathcal{N}$ will be called a null-class if

- (1) $S \subset S'$ and $S' \in \mathcal{N} \implies S \in \mathcal{N}$,
- (2) $\mathcal{N}$ is closed under countable unions.

A Borel null class is a null class such that any $S \in \mathcal{N}$ is contained in a Borel set $S' \in \mathcal{N}$.

Note that $\mu \ll \mu'$ if and only if $\mathcal{N}(\mu) \supset \mathcal{N}(\mu')$. In particular, two measures are mutually absolutely continuous if and only if they have the same null-class. Thus, although the measures μ_j are not canonically associated to A , the null-class of $\mu = \sum_j 2^{-j} \mu_j$ is.

More generally, we will say that a vector x has maximal spectral support for A if $m_{x,x}$ has the same null-class as $\sum_j 2^{-j} \mu_j$. That is, if $m_{x,x}(S) = 0$ implies $m_{x',x'}(S) = 0$ for all x' . We have already seen, in the proof of the corollary, that vectors with maximal spectral support exist. Zornifying we may prove

THEOREM 13.4. *There exists a spectral representation such that for each $j = 1, \dots, N$, $T_j 1$ has maximal spectral support for the restriction of A to $\oplus_{k=j}^N H_k$.*

In such a spectral representation, $\mu_{j+1} \ll \mu_j$ for each j , so $\mathcal{N}(\mu_{j+1}) \supset \mathcal{N}(\mu_j)$. Furthermore, the null-classes $\mathcal{N}(\mu_j)$, $j = 1, \dots, N$ are canonically associated to A . We can put this in a somewhat simpler form.

THEOREM 13.5 (Spectral theorem with multiplicities). *Let $A \in \mathcal{L}(H)$ be self-adjoint. Then there is a unique Borel null-class $\mathcal{N}$, and a Borel measurable function $\Xi : \sigma(A) \rightarrow \mathbb{N} \cup \{\infty\}$, unique up to redefinition on sets in $\mathcal{N}$, such that given any Borel measure μ on $\sigma(A)$ of class $\mathcal{N}$, and the associated Hilbert space*

$$L^2(\mu, \mathbb{C}^\Xi) = \{ \mathbf{f} \in L^2(\mu; \ell^2(\mathbb{N})) : \mathbf{f}(\lambda, n) = 0 \text{ if } n \geq \Xi(\lambda) \},$$

there is an isometry $T : L^2(\mu; \mathbb{C}^\Xi) \rightarrow H$ such that

$$AT = TM$$

with $M\mathbf{f}(\lambda, n) = \lambda \mathbf{f}(\lambda, n)$.

PROOF. Let μ_j be as above. Since $\mu_{j+1} \ll \mu_j$ for all j , we see that $\mu_j \ll \mu$, so there are non-negative $F_j \in L^1(\mu_1)$ such that $d\mu_j(\lambda) = F_j(\lambda)d\mu_1(\lambda)$. We set $F_1(\lambda) = 1$. We can choose F_j so that $F_j(\lambda) = 0 \implies F_{j+1}(\lambda) = 0$.

Let $\mathcal{N} = \mathcal{N}(\mu_1)$. We have seen that this null-class is canonical. Let

$$\Xi(\lambda) = \sup\{j : F_j(\lambda) > 0\}.$$

So $\Xi(\lambda) \geq 1$.

Now, let μ a Borel measure with null-class $\mathcal{N}$. Then $\mu \ll \mu_1$ so $d\mu = F d\mu_1$ with $F \in L^1(\mu_1)$. Since $\mu_1 \ll \mu$, we have $1/F \in L^1(\mu)$. We define a map $B : L^2(\mu; \ell^2(\mathbb{N})) \rightarrow \oplus_j L^2(\mu_j)$ as follows:

$$[B\mathbf{f}]_j(\lambda) = \begin{cases} \sqrt{\frac{F(\lambda)}{F_j(\lambda)}} \mathbf{f}(\lambda, j) & \text{if } F_j(\lambda) > 0. \\ 0 & \text{if } F_j(\lambda) = 0. \end{cases}$$

EXERCISE 28. Show that B is an isometry.

Now let $T : L^2(\mu; \ell^2(\mathbb{N})) \rightarrow H$ be

$$T\mathbf{f} = \sum_j T_j [B\mathbf{f}]_j.$$

So T is an isometry and $AT = TM$ as claimed.

EXERCISE 29. Suppose that Ξ and Ξ' are distinct multiplicity functions, so $\{\lambda : \Xi(\lambda) \neq \Xi'(\lambda)\} \notin \mathcal{N}$. Let M denote multiplication by λ on $L^2(\mu, \mathbb{C}^\Xi)$ and M' denote multiplication by λ on $L^2(\mu, \mathbb{C}^{\Xi'})$. Show that there is no isometry $T : L^2(\mu, \mathbb{C}^\Xi) \rightarrow L^2(\mu, \mathbb{C}^{\Xi'})$ such that $M'T = TM$. $\square$

The function Ξ is the multiplicity function.

EXERCISE 30. Show that if λ is an eigenvalue then $\Xi(\lambda)$ is the dimension of the corresponding eigenspace.

LECTURE 14

Jacobi matrix representations and orthogonal polynomials

This topic is not in the book, however I think it provides a very useful perspective on the subject. To start let us define:

DEFINITION 14.1. A vector $x \in H$ is cyclic for a self-adjoint operator $A \in \mathcal{L}(H)$ if $\{f(A)x : f \in C(\sigma(A))\}$ is dense in H .

Not every self-adjoint operator has a cyclic vector. Indeed it is easy to see that A has a cyclic vector if and only if the multiplicity of A is uniformly equal to 1. However, there is really no loss in studying operators with cyclic vectors. Indeed, the spectral theorem shows that every self-adjoint operator is a direct sum of operators possessing cyclic vectors.

THEOREM 14.1. *Let $A \in \mathcal{L}(H)$ be self-adjoint with a cyclic vector x . If H is infinite dimensional (and separable), then there is an isometry $T : H \rightarrow \ell^2(\mathbb{N})$ such that*

- (1) $Tx = \|x\| \delta_0$
- (2) $TAT^\dagger = J$, with J a tri-diagonal semi-infinite matrix, symmetric matrix with real entries on the diagonal and positive entries on the super diagonal.

REMARK. That is

$$J\mathbf{s}_n = \begin{cases} a_n\mathbf{s}_{n+1} + b_n\mathbf{s}_n + a_{n-1}\mathbf{s}_{n-1} & n \geq 1, \\ a_0\mathbf{s}_1 + b_0\mathbf{s}_0 & n = 0, \end{cases}$$

with $b_n \in \mathbb{R}$ and $a_n > 0$, corresponding to the semi-infinite matrix

$$J \sim \begin{pmatrix} b_0 & a_0 & & \\ a_0 & b_1 & a_1 & \\ & a_1 & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix}.$$

Such matrices are called Jacobi matrices.

PROOF. Assume without loss that $\|x\| = 1$. Let $\mu = m_{x,x}$. Then there is an isometry $S : H \rightarrow L^2(\mu)$ such that

$$Sx = 1, \quad \text{and} \quad SA = \lambda S.$$

Thus we may assume without loss of generality that $x = 1 \in L^2(\mu)$ and $Af(\lambda) = \lambda f(\lambda)$ with μ some compactly supported Borel measure on the real line.

Consider λ^n in $L^2(\mu)$. The span of this sequence is dense in L^2 . Furthermore, the sequence is linearly independent. Indeed, if this were not so then we would have

$$\lambda^n = \sum_{j=0}^{n-1} d_j \lambda^j$$

for some n , with some coefficients d_j . But then

$$\lambda^{n+1} = \sum_{j=0}^{n-1} d_j \lambda^{j+1} = d_{n-1} \lambda^n \sum_{j=1}^{n-1} d_{j-1} \lambda^j = \sum_{j=0}^{n-1} d_{n-1} d_j \lambda^j + \sum_{j=1}^{n-1} d_{j-1} \lambda^j \in \text{span}\{\lambda^j : j = 0, \dots, n-1\}.$$

Continuing by induction we would have λ^n in the span of $\{\lambda^j\}_{j=1}^{n-1}$. Then $H = L^2(\mu)$ would be finite dimensional, contrary to the hypothesis.

Now perform the Gram-Schmidt process on this sequence to produce a sequence of polynomials $\phi_n \in L^2(\mu)$. More specifically, let P_n denote projection of λ^n onto the space spanned by $1, \dots, \lambda^n$. Then

$$\phi_n = \frac{1}{\|\lambda^n - P_{n-1}\lambda^n\|} (\lambda^n - P_{n-1}\lambda^n).$$

That is ϕ_n is

- (1) normalized: $\|\phi_n\| = 1$,
- (2) non-negative for large λ , since the leading term is $\lambda^n / \|\lambda^n - P_{n-1}\lambda^n\| > 0$, and
- (3) orthogonal to any polynomial of degree $< n$.

(In fact, ϕ_n is the unique polynomial with these properties.) Thus ϕ_n is of the form

$$\phi_n(\lambda) = \sum_{j=0}^n c_j(n) \lambda^j$$

with coefficients $c_j(n)$ that satisfy

$$c_j(n) = \frac{1}{\|\lambda^n - P_{n-1}\lambda^n\|} m_j(n),$$

with $m_n(n) = 1$ and $\{m_j(n)\}_{j=0}^{n-1}$ is the minimizer for

$$F(\{\alpha_j\}_{j=0}^{n-1}) = \left\| \lambda^n - \sum_{j=0}^{n-1} \alpha_j \lambda^j \right\|^2 = \int \left| \lambda^n - \sum_{j=0}^{n-1} \alpha_j \lambda^j \right|^2 d\mu(\lambda).$$

The minimizer $\{m_j(n)\}$ is clearly real valued, so $c_j(n)$ are real.

The polynomials are orthonormal with respect to the weight μ :

$$\int \phi_n(\lambda) \phi_m(\lambda) d\mu(\lambda) = \delta_{n,m}.$$

Since 1 is cyclic, ϕ_n form a basis for $L^2(\mu)$. Let the isometry

$$T : L^2(\mu) \rightarrow \ell^2(\mathbb{N})$$

be expansion with respect to this basis:

$$[Tf]_n = \langle f, \phi_n \rangle.$$

What does A do to ϕ_n ? Well

$$A\phi_n(\lambda) = \lambda\phi_n(\lambda),$$

is a polynomial of degree $n+1$. If $m > n+1$ then

$$\langle \phi_m, A\phi_n \rangle = 0$$

because ϕ_m is orthogonal to all polynomials of degree less than m . Likewise, if $m < n - 1$ then

$$\langle \phi_m, A\phi_n \rangle = \langle A\phi_m, \phi_n \rangle = 0.$$

It follows that

$$A\phi_n = \langle A\phi_n, \phi_n \rangle \phi_n + \langle A\phi_n, \phi_{n+1} \rangle \phi_{n+1} + \langle A\phi_n, \phi_{n-1} \rangle \phi_{n-1}.$$

Let $b_n = \langle A\phi_n, \phi_n \rangle$ and $a_n = \langle A\phi_n, \phi_{n+1} \rangle$. Then this equation may be rewritten

$$A\phi_n = b_n \phi_n + a_n \phi_{n+1} + a_{n-1} \phi_{n-1},$$

with the convention that $a_{-1} = 0$.

Expanding the polynomials into sums of monomials gives

$$\sum_{j=0}^n c_j(n) \lambda^{j+1} = \sum_{j=0}^n b_n c_j(n) \lambda^j + \sum_{j=0}^{n+1} a_n c_j(n+1) \lambda^j + \sum_{j=0}^{n-1} a_{n-1} c_j(n-1) \lambda^j.$$

Thus,

$$\begin{aligned} a_n c_{n+1}(n+1) &= c_n(n), & a_n c_n(n+1) &= c_{n-1}(n) - b_n c_n(n), \\ a_n c_j(n+1) &= c_{j-1}(n) - b_j c_j(n) - a_{n-1} c_j(n-1) & \text{for } j = 1, \dots, n-1, \\ & \text{and } a_n c_0(n+1) &= -b_n c_0(n) - a_{n-1} c_0(n-1). \end{aligned}$$

In particular, since $c_n(n) > 0$ for all n we see that $a_n = c_n(n)/c_{n+1}(n+1) > 0$. $\square$

Thus we have the following correspondence: any compactly supported Borel measure μ on $\mathbb{R}$ gives rise to a tridiagonal matrix

$$J = J(\mu) = \begin{pmatrix} b_0 & a_0 & & \\ a_0 & b_1 & a_1 & \\ & a_1 & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix},$$

such that the solution $\phi_n(\lambda)$ to

$$J\phi = \lambda\phi, \quad \phi_0 = 1,$$

that is

$$\lambda\phi_n(\lambda) = b_n \phi_n(\lambda) + a_n \phi_{n+1}(\lambda) + a_{n-1} \phi_{n-1}(\lambda),$$

is the n^{th} orthonormal polynomial in $L^2(\mu)$. Conversely any Jacobi matrix, with $b_j \in \mathbb{R}$ and $a_j > 0$ bounded sequences gives rise to a self-adjoint operator on $\ell^2(\mathbb{N})$, and thus a measure $\mu = m_{x,x}$ with $x = \delta_0$. Thus

self adjoint operators with cyclic vectors

$\equiv$ compactly supported Borel measures on the real line

$\equiv$ Jacobi matrices.

LECTURE 15

Generalized eigenvectors for Jacobi matrices

If we have a measure μ and the associated Jacobi matrix J , then for fixed λ , the sequence $\phi(\lambda) = \{\phi_n(\lambda)\}_{n=0}^\infty$ is a solution to

$$J\phi(\lambda) = \lambda\phi(\lambda).$$

That is $\phi(\lambda)$ is an eigenvector! In the end, we have constructed eigenvectors for “any ” self-adjoint operator. However, we have constructed too many, since this holds for every real λ , even $\lambda \notin \sigma(A)$. Clearly, some of these eigenvectors have nothing to do with spectral theory.

EXERCISE 31. Let J be the Jacobi matrix with $b_n = 0$ and $a_n = 1$,

$$J = \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & 1 & \\ & 1 & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix}.$$

Show that $\|J\| \leq 2$ and that the eigenvector equation has solution

$$\phi_n(\lambda) = \frac{\sin(k(n+1))}{\sin(k)},$$

with k a solution of $\lambda = 2\cos(k)$. For $|\lambda| \leq 2$ the possible values of k are real and the resulting solution is bounded. For $\lambda > 2$, $k = i\kappa$ is purely imaginary and the solution,

$$\phi_n(\lambda) = \frac{\sinh(\kappa(n+1))}{\sinh(\kappa)},$$

with $\lambda = 2\cosh(\kappa)$ grows exponential as $n \rightarrow \infty$.

It is no accident that the solution grows exponentially as $n \rightarrow \infty$ for λ outside the spectrum of J :

THEOREM 15.1. *Let J be a Jacobi matrix and suppose $\lambda \notin \sigma(J)$ then*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln |\phi_n(\lambda)| > 0.$$

PROOF. Let $\phi_n = \phi_n(\lambda)$. Suppose, on the contrary, that $\limsup_{n \rightarrow \infty} \frac{1}{n} \ln |\phi_n| \leq 0$. Then for any $\mu > 0$ we have $e^{-\mu n} \phi_n \in \ell^2$.

Fix μ for the moment and let $\psi_n = e^{-\beta n} \phi_n$. Then

$$\lambda \psi_n = a_n e^\beta \psi_{n+1} + b_n \psi_n + a_{n-1} e^{-\beta} \psi_{n-1}.$$

That is ψ_n is an eigenvector of the matrix

$$J_\beta = J + \begin{pmatrix} 0 & (e^\beta - 1)a_0 & & \\ (e^{-\beta} - 1)a_0 & 0 & \ddots & \\ & & \ddots & \ddots \end{pmatrix}.$$

So, for small β ,

$$\|J_\beta - J\| = O(\beta).$$

It follows that, for small μ ,

$$\sigma(J_\beta) \subset \{\lambda' : \text{dist}(\lambda', \sigma(J)) \leq C\beta\}.$$

Hence if $\lambda \notin \sigma(J)$ we can tune μ to be small enough that $\lambda \notin \sigma(J_\beta)$. Thus ψ cannot be an ℓ^2 eigenvector and so the claimed condition must hold. $\square$

REMARK. The argument given here is known in mathematical physics as the “Combes-Thomas” argument. Note that J_β is not self adjoint.

On the flip side, we may control the growth of the polynomials $\phi_n(\lambda)$ for λ in a support of the spectral measure $\mu = m_{\delta_0, \delta_0}$.

THEOREM 15.2. *Fix $\epsilon > 0$. For μ almost every λ ,*

$$\frac{1}{(1+n)^{\frac{1}{2}+\epsilon}} \phi_n(\lambda) \in \ell^2.$$

PROOF. Note that $\frac{1}{(1+n)^{\frac{1}{2}+\epsilon}} \phi_n$ is the “Fourier expansion” of a function in $L^2(\mu)$, since

$$\left\| \frac{1}{(1+n)^{\frac{1}{2}+\epsilon}} \phi_n \right\|^2 = \sum_n \frac{1}{(1+n)^{1+\epsilon}} < \infty.$$

Since

$$\left\| \frac{1}{(1+n)^{\frac{1}{2}+\epsilon}} \phi_n \right\|^2 = \int \sum_n \frac{1}{(1+n)^{1+\epsilon}} |\phi_n(\lambda)|^2 d\mu(\lambda),$$

it follows from Fubini that the integrand on the r.h.s. is finite almost everywhere. $\square$

COROLLARY 15.3. *The spectrum of a Jacobi matrix is the closure of the set of real numbers such that there is a polynomially bounded solution to $J\mathbf{s} = \lambda\mathbf{s}$.*

So we have actually succeeded in producing an “eigenvector” expansion for the Jacobi matrix. The catch is that the eigenvectors need not lie in ℓ^2 — they may lie in a space of sequences with polynomial growth.

If we have a Hilbert space H and a self-adjoint A with a cyclic vector ϕ_0 we can pull back the “generalized eigenfunctions” in the Jacobi matrix representation as follows. Let $\phi_n \in H$ be the elements obtained by running the Gram-Schmidt process on $A^n \phi_0$. (If A is multiplication by λ on an L^2 space then ϕ_n are the orthonormal polynomials.) So $T\phi_n = \delta_n \in \ell^2(\mathbb{N})$. Let us introduce the following “scale of Hilbert spaces”

$$H_+ \subset H \subset H_-,$$

where

$$H_+ = \{x \in H : \sum_{n=0}^{\infty} (1+n)^{1+\epsilon} |\langle x, \phi_n \rangle|^2 < \infty\},$$

which is a Hilbert space in the inner product

$$\langle x, y \rangle_+ = \sum_{n=0}^{\infty} (1+n)^{1+\epsilon} \langle x, \phi_n \rangle \langle \phi_n, y \rangle.$$

The second space, H_- is the dual of H_+ , the space of linear functionals on H_+ . It is a Hilbert space under the inner product induced by the Riesz theorem identifying H_+ and H_- . However, if we ignore the Riesz theorem then we can embed H into H_- using the H inner product:

$$\ell_x(x_+) = \langle x_+, x \rangle, \quad x_+ \in H_+.$$

This defines a map $J : H \rightarrow H_-$. The Hilbert space H_- can also be thought of as the completion of H in the norm

$$\|x\|_- = \sup_{\|x_+\| \leq 1} |\langle x, x_+ \rangle|.$$

EXERCISE 32. Show that $\|x\|_-$ is a norm and that $\|x\|_-^2 = \sum_n \frac{1}{(1+n)^{1+\epsilon}} |\langle x, \phi_n \rangle|^2$.

Think of H_+ as a space of “smooth vectors” and H_- as a space of “distributions.” The point of the theorem is that eigenvectors may not lie in H , but the relevant ones are guaranteed to lie in a space of distributions. We will return to this notion later once we have introduced the notion of “trace class.” Then we will be able to develop a theory of generalized eigenvector expansions that makes no reference to Jacobi matrices.

LECTURE 16

Unbounded operators

Reading: Chapter 32

THEOREM 16.1 (Hellinger Toeplitz). *Suppose $A : H \rightarrow H$ is a linear operator defined on all of H and that*

$$\langle Ax, y \rangle = \langle x, Ay \rangle$$

for all $x, y \in H$. Then A is bounded.

PROOF. Suppose $x_n \rightarrow x$ and that Ax_n is convergent to $u \in H$. Then

$$\langle u, y \rangle = \lim_{n \rightarrow \infty} \langle Ax_n, y \rangle = \lim_{n \rightarrow \infty} \langle x_n, Ay \rangle = \langle x, Ay \rangle = \langle Ax, y \rangle$$

for all $y \in H$. Thus $u = Ax$. That is A is a closed operator. By the closed graph theorem, A is bounded. $\square$

The main point of this theorem is a negative result. If we are interested in an unbounded operator like $-\frac{d}{dx^2}$ on $L^2(\mathbb{R})$ which is formally symmetric on a dense subspace

$$\left\langle -\frac{d}{dx^2}u, v \right\rangle = \left\langle u, -\frac{d}{dx^2}v \right\rangle, \quad u, v \in C^2 \cap L^2$$

then there is no extension of $-\Delta$ to all of L^2 which is also symmetric. (Such an extension would necessarily be unbounded, since for instance,

$$-\frac{d}{dx^2}e^{-ikx}f(x) = -f''(x) + 2ikf'(x) + k^2f(x),$$

so

$$\left\| -\frac{d}{dx^2}e^{-ikx}f(x) \right\|_{L^2} \geq k^2 \|f\|_{L^2} - 2|k| \|f'\|_{L^2} - \|f''\|_{L^2} \rightarrow \infty, \quad k \rightarrow \infty,$$

although $\|e^{-ikx}f\|_{L^2} = \|f\|_{L^2}$ remains finite.)

Thus to talk about unbounded symmetric operators we must give up the notion that an operator is defined everywhere.

DEFINITION 16.1. A linear operator A on a Hilbert space H is a linear map $A : \mathcal{D} \rightarrow H$ with $\mathcal{D} = \mathcal{D}(A)$ a subspace of H . If $\mathcal{D}$ is dense in H , we say that A is densely defined and define the adjoint $A^\dagger$ of A to be the linear operator with

$$\mathcal{D}(A^\dagger) = \{v : |\langle Au, v \rangle| \leq C \|u\| \text{ for all } u \in \mathcal{D}(A)\}$$

such that

$$\langle Au, v \rangle = \langle u, A^\dagger v \rangle$$

for $u \in \mathcal{D}(A)$ and $v \in \mathcal{D}(A)$.

REMARK. By the Riesz theorem $A^\dagger v$ is uniquely defined on $\mathcal{D}(A^\dagger)$ since $\mathcal{D}(A)$ is dense. It may happen that $A^\dagger$ is not densely defined.

DEFINITION 16.2. A linear operator A is self adjoint if $A = A^\dagger$. That is, if $\mathcal{D}(A) = \mathcal{D}(A^\dagger)$ and

$$\langle Au, v \rangle = \langle u, Av \rangle \quad u, v \in \mathcal{D}(A).$$

Our goal in the coming days is to prove

THEOREM 16.2. Let A be a self-adjoint operator on a Hilbert space H , with domain $\mathcal{D}$. Then there is a projection valued measure E defined on Borel subsets of $\mathbb{R}$ such that

- (1) $E(\emptyset) = 0$, $E(\mathbb{R}) = I$
- (2) For any pair of measurable sets S, T , $E(S)E(T) = E(S \cap T)$
- (3) For every measurable set S , $E(S)^\dagger = E(S)$
- (4) E commutes with A . That is, for any measurable set S , $E(S)$ maps the domain $\mathcal{D}$ of A into $\mathcal{D}$, and for all u in $\mathcal{D}$, $AE(S)u = E(S)Au$.
- (5) $\mathcal{D} = \{u : \int t^2 d\langle E((-\infty, t))u, u \rangle < \infty\}$,
- (6) $Au = \int t dE((-\infty, t))u$

That is E is just like the projection valued measure for a bounded self adjoint operator, except it doesn't have compact support.

We will give several proofs of this theorem, because they expose different ideas. To begin we should define the resolvent set and resolvent of an unbounded operator.

DEFINITION 16.3. Let A be a linear operator. We say that a linear operator B is an extension of A if i.) $\mathcal{D}(B) \supset \mathcal{D}(A)$ and

$$Au = Bu \quad \text{for all } u \in \mathcal{D}(A),$$

denoted $A \subset B$. An operator A is symmetric if $A \subset A^\dagger$

Symmetric is not the same as self-adjoint.

EXERCISE 33. For instance, let $Au = -u''$ for $u \in \mathcal{D}(A) = C_0^2((0, 1)) \subset L^2(0, 1)$. Find $A^\dagger$ and show that A is symmetric but not self-adjoint.

In the last example A has a self-adjoint extension. However, it may happen that a symmetric operator has no self-adjoint extensions.

The sorts of linear operators we will deal with are closed in the following sense

DEFINITION 16.4. A linear operator A is closed if whenever $x_n \in \mathcal{D}(A)$ are such that $x_n \rightarrow x \in H$ and Ax_n converges we have $x \in \mathcal{D}(A)$ and $Ax_n \rightarrow Ax$.

EXERCISE 34. If A is densely defined, show that $A^\dagger$ is closed.

EXERCISE 35. An operator A is called closeable if it has a closed extension. The closed extension with smallest domain is called the closure of A , denoted $\overline{A}$. Show that densely defined A is closeable if and only if $A^\dagger$ is densely defined, in which case $A \subset (A^\dagger)^\dagger$ and use the previous exercise.

DEFINITION 16.5. Let A be a densely defined closed linear operator on a Hilbert space. The resolvent set $\rho(A)$ is the set of all $z \in \mathbb{C}$ such that $A - zI$ maps $\mathcal{D}(A)$ onto H in a one-to-one fashion. The resolvent is the map $(A - zI)^{-1} : H \rightarrow \mathcal{D}(A)$.

PROPOSITION 16.3. Let $z \in \rho(A)$ with A a densely defined closed linear operator. Then $(A - zI)^{-1}$ is bounded.

PROOF. By definition $(A - zI)^{-1}$ is everywhere defined. By the closed graph theorem, it suffices to show that $(A - zI)^{-1}$ is closed. Let $x_n \in H$ be such that $x_n \rightarrow x$ and $(A - zI)^{-1}x_n \rightarrow u$. Then

$$(A - zI)(A - zI)^{-1}x_n = x_n \rightarrow x,$$

so since $A - zI$ is closed, $u \in \mathcal{D}(A)$ and $(A - zI)u = x$. Thus $x = (A - zI)^{-1}u$. $\square$

We used:

EXERCISE 36. Let A be a closed linear operator and let $B \in \mathcal{L}(H)$ be bounded. Define $A + B$ on $\mathcal{D}(A)$ by $(A + B)u = Au + Bu$. Show that $A + B$ is closed.

THEOREM 16.4. *The resolvent set $\rho(A)$ of a densely defined closed linear operator is open, $(A - zI)^{-1}$ is a strongly analytic function of $z \in \rho(A)$, and the resolvent identity*

$$(A - zI)^{-1} - (A - wI)^{-1} = (w - z)(A - zI)^{-1}(A - wI)^{-1}$$

holds.

PROOF. This proceeds along the same lines as in Banach algebras. The key is the Neumann series expansion

$$(A - wI)^{-1} = \sum_{n=0}^{\infty} (w - z)^n (A - zI)^{-n-1},$$

from which it follows that $(A - wI)^{-1}$ is analytic in w in a disc of radius $1/\|(A - zI)^{-1}\|$ centered at z for any $z \in \rho(A)$. That $\rho(A)$ is open follows. The resolvent identity can be verified by explicit calculation since

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

on a dense subset. $\square$

DEFINITION 16.6. The spectrum $\sigma(A)$ of a densely defined closed linear operator A is the complement of the resolvent set $\rho(A)$.

THEOREM 16.5. *The spectrum $\sigma(A)$ of a densely defined closed linear operator A is a closed subset of $\mathbb{C}$. If A is self-adjoint then $\sigma(A) \subset \mathbb{R}$ and*

$$\|(A - zI)^{-1}\| \leq \frac{1}{\operatorname{Im} z}.$$

REMARK. We will see later that $\sigma(A) \neq \emptyset$ if A is self-adjoint. However there are densely defined closed operators with empty spectrum:

EXERCISE 37. Let $A = \frac{d}{dx}$ on the domain $\mathcal{D}$ of absolutely continuous functions f on $[0, 1]$ with $f(0) = 0$ and $f' \in L^2(0, 1)$. Show that A is densely defined, closed and that $\sigma(A) = \emptyset$. (Hint: A is the inverse of the Volterra integral operator $Vf(x) = \int_0^x f(y)dy$, which is a compact operator with no eigenvalues.)

As the exercise shows, empty spectrum is in some sense the same as $\sigma(A) = \{\infty\}$.

PROOF. The spectrum is the complement of an open set, so it is closed. Now suppose A is self-adjoint and $z \notin \mathbb{R}$. We must show that $z \in \rho(A)$. For $u \in \mathcal{D}(A)$ we have

$$\|u\| \|(A - zI)u\| \geq |\langle (A - zI)u, u \rangle| \geq |\operatorname{Im} \langle (A - zI)u, u \rangle| = |\operatorname{Im} z| \|u\|^2,$$

so

$$\|(A - zI)u\| \geq |\operatorname{Im} z| \|u\|.$$

It follows that $A - zI$ is one to one, that $\operatorname{ran}(A - zI)$ is closed, and that the inverse map $(A - zI)^{-1} : \operatorname{ran}(A - zI) \rightarrow H$ satisfies

$$\|(A - zI)^{-1}u\| \leq \frac{1}{|\operatorname{Im} z|} \|u\|.$$

It remains to show that $\operatorname{ran}(A - zI)$ is dense, as then it follows that $\operatorname{ran}(A - zI) = H$. So suppose $v \perp \operatorname{ran} A - zI$. Then

$$\langle v, Au \rangle = z \langle v, u \rangle$$

for all $u \in \mathcal{D}(A)$. It follows that $v \in \mathcal{D}(A^\dagger) = \mathcal{D}(A)$ and $Av = zv$. But then $\langle Av, v \rangle = z \|v\|^2$, so upon taking imaginary parts we find $0 = \operatorname{Im} z \|v\|^2$. Since $\operatorname{Im} z \neq 0$ we have $v = 0$. $\square$

LECTURE 17

Functional calculus fo self-adjoint operators

Let A be a self-adjoint operator. Then the resolvent $R(z) = (A - zI)^{-1}$ is an $\mathcal{L}(H)$ -valued analytic map on $\mathbb{C} \setminus \mathbb{R}$. Let $u \in H$ and consider

$$F_u(z) = \langle R(z)u, u \rangle.$$

So $F_u : \mathbb{C} \setminus \mathbb{R} \rightarrow \mathbb{C}$ is analytic. Note that

$$\operatorname{Im} F_u(z) = \frac{1}{2i} [\langle R(z)u, u \rangle - \langle u, R(z)u \rangle] = \frac{1}{2i} \langle (R(z) - R(z)^\dagger)u, u \rangle.$$

EXERCISE 38. Show that $R(z)^\dagger = R(\bar{z})$.

So by the resolvent identity

$$\operatorname{Im} F_u(z) = \frac{1}{2i} (z - \bar{z}) \langle R(z)^\dagger R(z)u, u \rangle = \operatorname{Im} z \|R(z)u\|^2.$$

Thus, if $\operatorname{Im} z > 0$ then $\operatorname{Im} F_u(z) > 0$. So we can think of F_u as an analytic self map of the upper half plane $\{\operatorname{Im} z > 0\}$.

Last term we showed

THEOREM 17.1 (Herglotz-Riesz). *Let F be an analytic function in the unit disk D such that $\operatorname{Re} F \geq 0$ in D . Then there is a unique non-negative, finite, Borel measure μ on ∂D such that*

$$F(z) = \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \mu(d\theta) + i \operatorname{Im} F(0).$$

Conversely every analytic function in the disk with positive real part can be written in this form.

As a corollary, using the mapping $\zeta \mapsto z = i \frac{1+\zeta}{1-\zeta}$ we also showed (see Lecture 21 from the notes for last term) that

THEOREM 17.2. *Let F be an analytic map from the upper half plane $\{z : \operatorname{Im} z > 0\}$ into itself. Then there is a unique non-negative Borel measure μ on $\mathbb{R}$ and a non-negative number $A \geq 0$ such that*

$$\int_{\mathbb{R}} \frac{1}{1+x^2} d\mu(x) < \infty$$

and

$$F(z) = Az + \operatorname{Re} F(i) + \int_{\mathbb{R}} \frac{1+xz}{x-z} \frac{1}{1+x^2} d\mu(x).$$

COROLLARY 17.3 (Nevanlinna). *Every analytic self map F of the upper half plane that satisfies*

$$\limsup_{y \rightarrow \infty} y |F(iy)| < \infty \tag{*}$$

is of the form

$$F(z) = \int_{\mathbb{R}} \frac{1}{x-z} d\mu(x), \quad (\star\star)$$

with μ a non-negative measure of finite mass

$$\mu(\mathbb{R}) = \lim_{y \rightarrow \infty} y \operatorname{Im} F(iy).$$

PROOF. The condition $(\star)$ guarantees that $A = 0$. Also, we have

$$y |F(iy)| \geq y \operatorname{Im} F(iy) = \int_{\mathbb{R}} \operatorname{Im} \frac{1+ixy}{x-iy} \frac{y}{1+x^2} d\mu(x),$$

and

$$\operatorname{Im} \frac{1+ixy}{x-iy} = y \frac{1+x^2}{x^2+y^2}.$$

So

$$y |F(iy)| \geq \int_{\mathbb{R}} \frac{y^2}{x^2+y^2} d\mu(y).$$

By dominated convergence, the r.h.s. converges to $\mu(\mathbb{R})$ as $y \rightarrow \infty$. Thus $(\star)$ implies that $\mu(\mathbb{R}) < \infty$.

Finally, we have

$$\frac{1+xz}{x-z} \frac{1}{1+x^2} = \frac{1}{x-z} - \operatorname{Re} \frac{1}{x-i}.$$

So,

$$F(z) = \int_{\mathbb{R}} \frac{1}{x-z} d\mu(x) + \left[\operatorname{Re} F(i) - \operatorname{Re} \int_{\mathbb{R}} \frac{1}{x-i} d\mu(x) \right].$$

Since $F(iy) \rightarrow 0$ as $y \rightarrow \infty$, we conclude that the term in brackets on the r.h.s. is zero and the representation $(\star\star)$ holds. $\square$

Returning to the self-adjoint operator A , let us check condition $(\star)$ for $F_u(z) = \langle R(z)u, u \rangle$. By the theorem of last time,

$$y |F_u(iy)| \leq y \frac{1}{y} \|u\|^2 \leq \|u\|^2,$$

so $(\star)$ holds. We conclude the existence of a finite Borel measure $\mu_{u,u}$ such that

$$\langle R(z)u, u \rangle = \int_{\mathbb{R}} \frac{1}{\lambda-z} d\mu_{u,u}(\lambda).$$

This looks a lot like a spectral representation.

THEOREM 17.4. *Let A be a self-adjoint operator on a Hilbert space $\mathcal{H}$. To each pair of vectors $u, v \in \mathcal{H}$ there is associated a unique complex Borel measure of finite total variation on the real line, $\mu_{u,v}$, such that*

$$\langle (A - zI)^{-1}u, v \rangle = \int_{\mathbb{R}} \frac{1}{\lambda-z} d\mu_{u,v}(\lambda).$$

Furthermore

- (1) For each u , $\mu_{u,u}$ is non-negative with $\mu_{u,u}(\mathbb{R}) = \|u\|^2$.
- (2) The map $(u, v) \mapsto \mu_{u,v}$ is linear in u , skew linear in v , skew-symmetric $\overline{\mu_{u,v}} = \mu_{v,u}$, and the total variation of $\mu_{u,v}$ is bounded by $\|u\| \|v\|$.

PROOF. We have already obtained the non-negative measures $\mu_{u,u}$. The identity for the norm is equivalent to

$$\lim_{y \rightarrow \infty} y \operatorname{Im} \langle (A - iy)^{-1} u, u \rangle = \|u\|^2.$$

Since

$$y(A - iy)^{-1} = iI - iA(A - iy)^{-1} \quad (\star)$$

it suffices to show that $A(A - iy)^{-1}u \rightarrow 0$ as $y \rightarrow \infty$. Referring to the identity $(\star)$ we see that $\|A(A - iy)^{-1}\|$ is uniformly bounded. Thus it suffices to show $A(A - iy)^{-1}u \rightarrow 0$ as $y \rightarrow \infty$ for u in the dense set $\mathcal{D}(A)$. However for such u we then have

$$\|A(A - iy)^{-1}u\| = \|(A - iy)^{-1}Au\| \leq \frac{1}{y} \|Au\| \rightarrow 0.$$

We may define the measure $\mu_{u,v}$ by polarization,

$$\mu_{u,v} = \frac{1}{4} [\mu_{u+v,u+v} - \mu_{u-v,u-v} + i\mu_{u+iv,u+iv} - i\mu_{u-iv,u-iv}].$$

It is an easy exercise to see that the identity for the resolvent holds. Skew-linearity and skew-symmetry of the map is clear from the representation formula for the resolvent. To see the inequality for norms, note that $(u, v) \mapsto \int_S d\mu_{u,v}$ is a (possibly degenerate) inner product on $\mathcal{H}$. Thus by Cauchy-Schwarz we have

$$|\mu_{u,v}(S)| \leq \sqrt{\mu_{u,u}(S)\mu_{v,v}(S)} \leq \frac{1}{2} [\mu_{u,u}(S) + \mu_{v,v}(S)].$$

It follows that the total variation is bounded by

$$\|\mu_{u,v}\| \leq \frac{1}{2} [\mu_{u,u}(\mathbb{R}) + \mu_{v,v}(\mathbb{R})] = \frac{1}{2} [\|u\|^2 + \|v\|^2],$$

which is the *wrong inequality!* But this is not a problem as the inequality we have derived does not scale properly. What we have really learned is that if $\|u\|, \|v\| \leq 1$ then

$$\|\mu_{u,v}\| \leq 1.$$

The result follows by scaling. □

Differentiating each side of the identity for the resolvent we see that

$$\langle (A - zI)^{-n} u, v \rangle = (-1)^n \frac{d^{n-1}}{dz^{n-1}} \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{u,v}(\lambda) = \int_{\mathbb{R}} \frac{1}{(\lambda - z)^n} d\mu_{u,v}(\lambda).$$

It follows that any of

$$\left\langle \sum_{j=1}^n c_j (A - zI)^{-m_j} u, v \right\rangle = \int_{\mathbb{R}} \sum_{j=1}^n c_j (\lambda - z)^{m_j} d\mu_{u,v}(\lambda).$$

EXERCISE 39. Show that functions of the form $f(\lambda) = \sum_{j=1}^n c_j (\lambda - z)^{m_j}$ are dense in $C_0(\mathbb{R})$, the continuous functions on $\mathbb{R}$ that vanish at ∞ .

So we have a continuous functional calculus.

$$f \mapsto \langle f(A)u, v \rangle,$$

with $\langle f(A)u, v \rangle$ defined as a limit $\lim_{n \rightarrow \infty} \langle f_n(A)u, v \rangle$ where f_n is of the above form.

LECTURE 18

Spectral representation of self-adjoint operators

Last time we constructed a continuous functional calculus $f \mapsto f(A)$ from $C_0(\mathbb{R}) \rightarrow \mathcal{L}(H)$ for a self-adjoint operator A . We got this from the spectral measures $\mu_{u,v}$ constructed from the Riesz-Herglotz representation for the resolvent. Now consider the map which associates to any Borel set S in $\mathbb{R}$ the quadratic form

$$Q_S(u, v) = \int_S d\mu_{u,v}.$$

We have

- (1) $|Q_S(u, v)| \leq \|u\| \|v\|$
- (2) $Q_S(\bullet, \bullet)$ is a sesquilinear skew-symmetric form.

We have seen in Lecture 11 that any such form is the quadratic form of a bounded self-adjoint operator. So we have $E(S) \in \mathcal{L}(H)$ such that

$$\langle E(S)u, v \rangle = \mu_{u,v}(S).$$

Clearly we have

$$\langle f(A)u, v \rangle = \int_{\mathbb{R}} f(\lambda) d\langle E(-\infty, \lambda)u, v \rangle, \quad (\star)$$

for $f \in C_0(\mathbb{R})$. As you have guessed, $S \mapsto E(S)$ is the projection valued measure for A . So we have the theorem claimed in Lecture 16:

THEOREM. *Let A be a self-adjoint operator on a Hilbert space H , with domain $\mathcal{D}$. Then there is a projection valued measure E defined on Borel subsets of $\mathbb{R}$ such that*

- (1) $E(\emptyset) = 0, E(\mathbb{R}) = I$
- (2) For any pair of measurable sets S, T , $E(S)E(T) = E(S \cap T)$
- (3) For every measurable set S , $E(S)^\dagger = E(S)$
- (4) E commutes with A . That is, for any measurable set S , $E(S)$ maps the domain $\mathcal{D}$ of A into $\mathcal{D}$, and for all u in $\mathcal{D}$, $AE(S)u = E(S)Au$.
- (5) $\mathcal{D} = \{u : \int t^2 d\langle E((-\infty, t))u, u \rangle < \infty\}$,
- (6) $Au = \int t dE((-\infty, t))u$

PROOF. Given the functional calculus, the proof of (1), (2), and (3) is essentially the same as the corresponding result for bounded self-adjoint operators. (See Lecture 11.)

To show (4), let us first show that $E(S)f(A) = f(A)E(S)$ for any $f \in C_0(A)$. We have

$$\begin{aligned} \langle E(S)f(A)u, v \rangle &= \int_{\mathbb{R}} f(\lambda) d\langle E(-\infty, \lambda)u, E(S)v \rangle = \int_{\mathbb{R}} f(\lambda) d\langle E(S)E(-\infty, \lambda)u, v \rangle \\ &= \int_{\mathbb{R}} f(\lambda) d\langle E(-\infty, \lambda)E(S)u, v \rangle = \langle f(A)E(S)u, v \rangle \end{aligned}$$

by (2). In particular $(A - zI)^{-1}$ commutes with $E(S)$ for any $z \in \mathbb{C} \setminus \mathbb{R}$, so $E(S) : \mathcal{D} \rightarrow \mathcal{D}$ (recall that $\mathcal{D} = \text{ran}(A - zI)^{-1}$.) If $u \in \mathcal{D}$ then

$$\begin{aligned} E(S)Au &= (A - zI)(A - zI)^{-1}E(S)Au = (A - zI)E(S)(A - zI)^{-1}Au \\ &= (A - zI)E(S)u + z(A - zI)E(S)(A - zI)^{-1}u \\ &= AE(S)u - zE(S)u + zE(S)u = AE(S)u. \end{aligned}$$

To prove (5), note that if $v \in \mathcal{D}$ then $v = (A - iI)^{-1}u$ for some $u \in H$. So

$$\begin{aligned} \langle E(T)v, v \rangle &= \langle E(T)R(i)u, R(i)u \rangle \\ &= \langle R(-i)R(i)E(T)u, u \rangle = \frac{1}{2i} \langle [R(i) - R(-i)]E(T)u, u \rangle \\ &= \frac{1}{2i} \int_{\mathbb{R}} \left[\frac{1}{t - i} - \frac{1}{t + i} \right] d \langle E(-\infty, t)E(T)u, u \rangle \\ &= \int_{\mathbb{R}} \frac{1}{1 + t^2} d \langle E(-\infty, t)u, u \rangle. \end{aligned}$$

Thus $d \langle E(-\infty, t)v, v \rangle = \frac{1}{1+t^2} d \langle E(-\infty, t)u, u \rangle$. Thus

$$(1 + t^2) d \langle E(-\infty, t)v, v \rangle = d \langle E(-\infty, t)u, u \rangle$$

is a finite measure, so $\int_{\mathbb{R}} t^2 d \langle E(-\infty, t)v, v \rangle < \infty$ as claimed.

Conversely, suppose $\int_{\mathbb{R}} t^2 d \langle E(-\infty, t)v, v \rangle < \infty$. Given $u \in H$, since

$$|\langle E(S)v, u \rangle| \leq \|u\|^2 \|E(S)v\|^2 = \langle E(S)v, v \rangle,$$

the measure $d \langle E(-\infty, t)v, u \rangle$ is absolutely continuous with respect to $d \langle E(-\infty, t)v, v \rangle$. Thus

$$d \langle E(-\infty, t)v, u \rangle = F_u(t) d \langle E(-\infty, t)v, v \rangle,$$

and

$$\begin{aligned} \left| \int_{\mathbb{R}} t d \langle E(-\infty, t)v, u \rangle \right| &\leq \left[\int_{\mathbb{R}} t^2 d \langle E(-\infty, t)v, v \rangle \int_{\mathbb{R}} |F_u(t)|^2 d \langle E(-\infty, t)v, v \rangle \right]^{\frac{1}{2}} \\ &\leq \|u\| \|v\| \left[\int_{\mathbb{R}} t^2 d \langle E(-\infty, t)v, v \rangle \right]^{\frac{1}{2}}. \end{aligned}$$

Thus $u \mapsto \int_{\mathbb{R}} t d \langle E(-\infty, t)v, u \rangle$ is a bounded linear functional. So $\exists \tilde{v}$ such that

$$\int_{\mathbb{R}} t d \langle E(-\infty, t)v, u \rangle = \langle \tilde{v}, u \rangle.$$

Now consider $(A - iI)^{-1}\tilde{v}$. We have

$$\langle (A - iI)^{-1}\tilde{v}, u \rangle = \int_{\mathbb{R}} \frac{1}{t - i} d \langle E(-\infty, t)\tilde{v}, u \rangle,$$

and

$$\langle E(-\infty, t)\tilde{v}, u \rangle = \int_{\mathbb{R}} s d \langle E(-\infty, s)v, E(-\infty, t)u \rangle = \int_{(-\infty, t)} s d \langle E(-\infty, s)v, u \rangle.$$

So

$$d \int_{(-\infty, t)} s d \langle E(-\infty, s)v, u \rangle = t d \langle E(-\infty, t)v, u \rangle,$$

and

$$\langle (A - iI)^{-1}\tilde{v}, u \rangle = \int_{\mathbb{R}} \frac{t}{t - i} d\langle E(-\infty, t)v, u \rangle = \langle v, u \rangle + i\langle (A - iI)^{-1}v, u \rangle.$$

We conclude that

$$(A - iI)^{-1}\tilde{v} = v + i(A - iI)^{-1}v.$$

It follows that $v \in \mathcal{D}(A)$ and

$$Av = A(A - iI)^{-1}(\tilde{v} - v) = \tilde{v} - v + i(A - iI)^{-1}(\tilde{v} - v) = \tilde{v}.$$

In particular, we have proved that for any $v \in \mathcal{D}(A)$

$$\langle Av, u \rangle = \int_{\mathbb{R}} t d\langle E(-\infty, t)v, u \rangle \quad \forall u.$$

EXERCISE 40. Show that

$$Av = \int_{\mathbb{R}} t dE(-\infty, t)v$$

as an improper norm convergent Stieltjes integral. That is, show that

$$Av = \lim_{r \rightarrow \infty} \int_{-r}^r t dE(-\infty, t)v,$$

where the integrals $\int_{-r}^r t dE(-\infty, t)v$ can be understood as Riemann-Stieltjes integrals, that is as a limit of Riemann sums

$$\sum_{j=1}^n t_j E(I_j)v,$$

where $t_j \in I_j$ and I_j is a partition of $(-r, r)$ into disjoint intervals. □

As in the bounded case we have

THEOREM 18.1. *Let A be a self-adjoint operator on a Hilbert space H . Then there is a finite measure space M, μ and an isometry $T : H \rightarrow L^2(M)$ such that for every $u \in \mathcal{D}(A)$*

$$[TAu](m) = \Phi(m)Tu(m),$$

where $\Phi(m)$ is a real valued measurable function on M . If A is unbounded, then the function Φ is unbounded and

$$T\mathcal{D}(A) = \{f \in L^2(m) : \Phi f \in L^2(m)\}.$$

Spectral multiplicities and the decomposition into point, singular continuous and absolutely continuous spectrum can be handled just as in the bounded case.

LECTURE 19

Unitary operators and von Neumann's proof the spectral theorem; Positive operators and the polar decomposition

Unitary operators.

THEOREM 19.1. *Let $U \in \mathcal{L}(H)$ be a unitary operator ($U^\dagger U = UU^\dagger = I$). Then $\sigma(U) \subset \{|z| = 1\} =: S_1$ and there is a projection valued measure E on Borel subsets of S_1 such that*

$$U = \int_{S_1} e^{i\theta} dE(\theta).$$

REMARK. There is a functional calculus for U given by

$$f(U) = \int_{S_1} f(\theta) dE(\theta), \quad f \in C(S_1).$$

PROOF. Since

$$\|U\| = \sup_{\|x\|=1} \|Ux\| = \sup_{\|x\|=1} \|x\| = 1,$$

we see that $\sigma(U) \subset \{|z| \leq 1\}$. Also, for $|z| < 1$, $(I - zU^\dagger)$ is invertible and

$$(U - z)U^\dagger(I - zU^\dagger)^{-1} = (I - zU^\dagger)(I - zU^\dagger)^{-1} = I = (I - zU^\dagger)^{-1}(I - zU^\dagger),$$

so $z \notin \sigma(U)$. Thus $\sigma(U) \subset S_1$. (In fact, we proved this previously in the context of C^* algebras.)

Now, construct a functional calculus for U by taking the norm limit of polynomials $p(U, U^\dagger)$. By Stone-Weierstrass, we learn to evaluate $f(U)$ for any $f \in C(S_1)$. By Riesz-Kakutani we find spectral measures $\mu_{x,y}$ for any $x, y \in H$ such that

$$\langle f(U)x, y \rangle = \int_{S_1} f(z) d\mu_{x,y}(z).$$

The PVM E is constructed from the spectral measures just as for self-adjoint operators. $\square$

THEOREM 19.2. *Let A be self-adjoint with domain $\mathcal{D} \subset H$. Then $U = (A - iI)(A + iI)^{-1}$ is a unitary map.*

PROOF. Since $A - iI$ maps $\mathcal{D}$ onto H in a one to one fashion and $(A + iI)^{-1}$ maps H onto $\mathcal{D}$ in a one to one fashion, we see that $U : H \rightarrow H$ is one to one and onto. Thus, it suffices to show that U is isometry.

Let $u \in H$ and let

$$v = (A + iI)^{-1}u, \quad w = Uu.$$

Then $v \in \mathcal{D}$ and

$$(A + iI)v = u, \quad (A - iI)v = w.$$

Thus

$$\|u\|^2 = \langle (A + iI)v, (A + iI)v \rangle = \|Av\|^2 - i\langle Av, v \rangle + i\langle v, Av \rangle + \|v\|^2 = \|Av\|^2 + \|v\|^2,$$

and likewise

$$\|w\|^2 = \langle (A - iI)v, (A - iI)v \rangle = \|Av\|^2 + \|v\|^2,$$

so $\|u\| = \|w\|$. □

The operator $U = (A - iI)(A + iI)^{-1}$ is called the “Cayley Transform” of A . Note that it is the fractional linear transformation

$$\Phi(\lambda) = \frac{\lambda - iI}{\lambda + iI}$$

evaluated at $\lambda = A$. This map maps $\mathbb{R}$ to S_1 with $\infty \mapsto 1$ and the $\{\operatorname{Im} z > 0\} \mapsto \{|z| < 1\}$. The original proof of the spectral Theorem for self-adjoint operators given by von Neumann appealed to the spectral theory of U . Specifically, if E_U is the PVM for U then we define

$$E(S) = E_U(\Phi(S)).$$

Positive operators and the polar decomposition.

DEFINITION 19.1. An unbounded operator A is positive (or form positive) if

$$\langle Av, v \rangle \geq 0$$

for all $v \in \mathcal{D}(A)$.

REMARK. An unbounded positive operator A may not be self-adjoint. We will see next time, however, that such an operator always has a positive self-adjoint extension.

THEOREM 19.3. *A self-adjoint operator A is positive if and only if $\sigma(A) \subset [0, \infty)$.*

PROOF. By the spectral theorem we may as well suppose that $H = L^2(M)$ and $Af(m) = \Phi(m)f(m)$ for some unbounded real valued function Φ on M . Clearly

$$\sigma(A) = \text{essential closure of } \{\Phi(m) : m \in M\},$$

where

essential closure of S = smallest closed set F such that $F^c \cap S$ has measure 0.

Thus $\sigma(A) \subset [0, \infty)$ if and only if $\Phi(m) \geq 0$ for almost every m . This latter condition is clearly necessary and sufficient for

$$\int_M \Phi(m) |f(m)|^2 d\mu(m) \geq 0$$

for all $f \in L^2(\mu)$. □

THEOREM 19.4. *Let A be self-adjoint and positive. Then there is a unique positive self-adjoint operator $\sqrt{A}$ such that $(\sqrt{A})^2 = A$.*

PROOF. Assume without loss that $H = L^2(M)$ and $Af(m) = \Phi(m)f(m)$ with $\Phi(m) \geq 0$. Then define $\sqrt{A}f(m) = \sqrt{\Phi(m)}f(m)$ on the domain

$$\mathcal{D}(\sqrt{A}) = \left\{ f : \sqrt{\Phi(m)}f(m) \in L^2 \right\}.$$

Clearly $\sqrt{A}$ is positive and self-adjoint.

EXERCISE 41. Verify that $\sqrt{A}$ is unique. □

REMARK. Note that we can define $F(A)$ for any function $F \in C(\sigma(A))$ by

$$F(A)u = \int_{\mathbb{R}} F(\lambda) dE(-\infty, \lambda)u,$$

on the domain

$$\left\{ u : \int_{\mathbb{R}} |F(\lambda)|^2 d\|E(-\infty, \lambda)u\|^2 < \infty \right\}.$$

THEOREM 19.5 (Polar decomposition). *Let A be a densely defined closed operator on a Hilbert space H . Then there is a positive self-adjoint operator $|A|$ with $\mathcal{D}(|A|) = \mathcal{D}(A)$ and a partial isometry $T : H \rightarrow H$ such that*

$$A = T|A|.$$

PROOF. Since A is closed, its domain $\mathcal{D}$ is a Hilbert space under the inner product

$$\langle u, v \rangle_1 = \langle Au, Av \rangle + \langle u, v \rangle.$$

Let $w \in H$ and note that

$$|\langle u, w \rangle| \leq \|u\| \|w\| \leq \|w\| \|u\|_1, \quad \text{for all } u \in \mathcal{D}.$$

By the Riesz theorem, there is a unique element $v \in \mathcal{D}$ such that

$$\langle u, w \rangle = \langle Au, Av \rangle + \langle u, v \rangle, \quad \text{for all } u \in \mathcal{D}.$$

Let $J : H \rightarrow \mathcal{D}$ denote the map $Jw = v$. It is well defined by uniqueness of v and is one-to-one since

$$\langle u, w \rangle = 0 \text{ for all } u \in \mathcal{D} \implies w = 0$$

since $\mathcal{D}$ is dense. Let $\mathcal{D}' = \text{ran } J \subset \mathcal{D}$ and define $M : \mathcal{D}' \rightarrow H$ to be

$$Mu = J^{-1}u - u.$$

(Formally, $J = (A^\dagger A + I)^{-1}$ so $M = A^\dagger A$. All of the “nonsense” is to make sense of this definition.)

EXERCISE 42. Show that M is positive and self-adjoint.

Since M is positive and self-adjoint, it has a square root. Let $|A| = \sqrt{M}$. Note that for $u \in \mathcal{D}'$

$$\|Au\|^2 = \|u\|_1^2 - \|u\|^2 = \langle u, J^{-1}u \rangle - \langle u, u \rangle = \langle u, Mu \rangle = \| |A|u \|^2.$$

Since $\mathcal{D}'$ is dense in $\mathcal{D}$ it follows that $\mathcal{D} = \mathcal{D}(|A|)$ and that

$$\|Au\| = \| |A|u \| \quad \text{for all } u \in \mathcal{D}.$$

Now let $T|A|u = Au$. This definition is ok because $Au = 0 \implies |A|u = 0$. Clearly T is an isometry from $\text{ran } |A|$ to $\text{ran } A$. Extend it (uniquely) to an isometry from $\overline{\text{ran } |A|} \rightarrow \overline{\text{ran } A}$ and then as you like to $\text{ran } |A|^\perp$. $\square$

LECTURE 20

Examples of self-adjoint extensions

In exercise 35 of Lecture 16 (exercise 21 of the homework), you showed that a densely defined operator A is closeable, i.e., has a closed extension, if and only if $A^\dagger$ is densely defined. In fact

THEOREM 20.1. *Let A be densely defined with densely defined adjoint $A^\dagger$. If A is closed then $(A^\dagger)^\dagger = A$. If A is only closeable then its closure $\overline{A} = (A^\dagger)^\dagger$.*

The proof of is based on

LEMMA 20.2. *Let A be densely defined, with graph $\Gamma(A) = \{(u, Au) : u \in \mathcal{D}(A)\}$. Then the graph of $A^\dagger$, $\Gamma(A^\dagger) = \{(u, A^\dagger u) : u \in \mathcal{D}(A^\dagger)\}$ is*

$$\Gamma(A^\dagger) = J(\Gamma(A))^\perp,$$

where $J(u, v) = (-v, u)$ and $\Gamma(A)^\perp$ is the annihilator of $\Gamma(A)$ in $H \oplus H$:

$$\Gamma(A)^\perp = \{(u, v) \in H \oplus H : \langle u, u' \rangle + \langle v, v' \rangle = 0 \text{ for all } (u', v') \in \Gamma(A)\}.$$

EXERCISE 43. Prove the lemma.

PROOF OF THEOREM. First suppose A is closed. Then $\Gamma(A)$ is a closed subspace of $H \oplus H$. Thus $[\Gamma(A)^\perp]^\perp = \Gamma(A)$. It follows that

$$\Gamma((A^\dagger)^\dagger) = J\Gamma(A^\dagger)^\perp = J^2\Gamma(A) = \Gamma(A),$$

so $A = (A^\dagger)^\dagger$.

EXERCISE 44. Verify that $\overline{A} = (A^\dagger)^\dagger$ if A is closeable. □

If A is symmetric then $A^\dagger \supset A$. It follows that a densely defined symmetric operator A is closeable.

DEFINITION 20.1. A densely defined symmetric operator A is essentially self adjoint if $\overline{A}$ is self-adjoint.

A useful criterion is

PROPOSITION 20.3. *Let A be densely defined and symmetric. Then A is essentially self-adjoint if and only if $A^\dagger$ is symmetric.*

PROOF. Since A is symmetric

$$A \subset \overline{A} \subset A^\dagger.$$

Now, for any densely defined closeable operator A we have $\overline{A}^\dagger = A^\dagger$. Thus, if A is essentially self-adjoint then $\overline{A} = A^\dagger$ is self-adjoint, hence symmetric. On the other hand if $A^\dagger$ is symmetric, so

$$A^\dagger \subset (A^\dagger)^\dagger = \overline{A},$$

we conclude that

$$A^\dagger = \overline{A} = \overline{A}^\dagger,$$

so $\overline{A}$ is self-adjoint. □

Examples.

PROPOSITION 20.4. *Let $Af(x) = i\frac{d}{dx}f(x)$ on $\mathcal{D} = C_c^\infty(\mathbb{R})$. Then A is self-adjoint.*

PROOF. By integration by parts, A is symmetric. Furthermore

$$\mathcal{D}(A^\dagger) = \left\{ f \in L^2 : \left| \int_{\mathbb{R}} f(x) \frac{d}{dx} u(x) dx \right| \leq c \sqrt{\int_{\mathbb{R}} |u(x)|^2 dx} \quad u \in C_c^\infty(\mathbb{R}) \right\}.$$

By the Riesz Frechet theorem, since C_c^∞ is dense, this is exactly the set of $f \in L^2$ such that the distributional derivative of $f \in L^2$:

$$\mathcal{D}(A^\dagger) = \left\{ f \in L^2 : \frac{d}{dx} f \in L^2 \right\}.$$

Furthermore it follows that $A^\dagger f(x) = i\frac{d}{dx}f(x)$ for all such f .

EXERCISE 45. Verify that integration by parts holds for all $f, g \in L^2$ such that the distributional derivatives $f', g' \in L^2$:

$$\int_{\mathbb{R}} f(x)g'(x)dx = - \int_{\mathbb{R}} f'(x)g(x)dx.$$

Thus $A^\dagger$ is symmetric and hence A is essentially self-adjoint. □

However, not every symmetric operator is essentially self-adjoint.

PROPOSITION 20.5. *Let $Af(x) = i\frac{d}{dx}f(x)$ on $\mathcal{D} = \{f \in C^1([0, 1]) : f(0) = f(1) = 0\} \subset L^2(0, 1)$. Then A is not essentially self-adjoint.*

PROOF. By IBP, A is symmetric:

$$\langle Af, g \rangle = \int_0^1 if'(x)\overline{g(x)}dx = i[f(1)g(1) - f(0)g(0)] + \int_0^1 f(x)\overline{ig'(x)}dx = \langle f, Ag \rangle,$$

since the boundary terms vanish if $f, g \in \mathcal{D}$.

EXERCISE 46. Show that $\mathcal{D}(A^\dagger) = \{f \in L^2(0, 1) : f' \in L^2\}$ and $A^\dagger f(x) = if'(x)$.

Now $A^\dagger$ is certainly not self-adjoint since if $f(x) = e^{-izx}$

$$A^\dagger f(x) = zf(x),$$

so every complex number $z \in \sigma(A^\dagger)$.

The closure $\overline{A}f(x) = if'(x)$ on

$$\mathcal{D}(\overline{A}) = \{f \in L^2(0, 1) : f' \in L^2 \text{ and } f(0) = f(1) = 0\}.$$

Note that $f' \in L^2 \implies f$ is continuous, so $f(0) = f(1) = 0$ makes sense. □

The operator A of the previous example does have self-adjoint extensions. In fact,

PROPOSITION 20.6. Let $\alpha \in [0, 2\pi)$ and let $A_\alpha f(x) = i \frac{d}{dx} f(x)$ on $\mathcal{D}_\alpha = \{f \in L^2(0, 1) : f' \in L^2 \text{ and } f(1) = e^{i\alpha} f(0)\}$. Then A_α is self-adjoint, and

$$\sigma(A_\alpha) = \{\alpha + 2\pi n : n \in \mathbb{N}\},$$

Each point $\alpha + 2\pi n \in \sigma(A_\alpha)$ is a non-degenerate eigenvalue. The eigenfunctions for A_α are an orthonormal basis for $L^2(0, 1)$.

PROOF.

EXERCISE 47. Show that A_α is self-adjoint.

EXERCISE 48. Show that λ is an eigenvalue of A_α if and only if $\lambda = \alpha + 2\pi n$ for some n . Conclude that $\{\alpha + 2\pi n\} \subset \sigma(A_\alpha)$.

EXERCISE 49. Fix $g \in L^2(0, 1)$. Obtain an explicit solution $f \in \mathcal{D}_\alpha$ to the equation

$$(A_\alpha - \alpha - \pi)f = g$$

in terms of an integral operator acting on g . (Hint: invert the derivative using an integrating factor. Use the constant of integration to satisfy the boundary condition required so that $f \in \mathcal{D}_\alpha$.)

EXERCISE 50. Show that the integral operator you obtained in the last exercise is compact and self-adjoint. Conclude that $\{\alpha + 2\pi n\} = \sigma(A_\alpha)$ and that the eigenfunctions you found are an orthonormal basis. $\square$

However, there are symmetric operators with no self-adjoint extensions

PROPOSITION 20.7. Let $Af(x) = i \frac{d}{dx} f(x)$ on $\{f \in L^2(0, \infty) : f' \in L^2 \text{ and } f(0) = 0\}$. Then A is closed and symmetric but has no self-adjoint extensions.

PROOF.

EXERCISE 51. Prove that A is closed and symmetric.

EXERCISE 52. Show that $A^\dagger f = if'$ on

$$\mathcal{D}(A^\dagger) = \{f \in L^2(0, \infty) : f' \in L^2\}.$$

Now $A^\dagger$ is clearly not symmetric since $A^\dagger e^{-x} = -ie^{-x}$. Note that

$$\dim \ker(A^\dagger - z) = \begin{cases} 0 & \operatorname{Im} z > 0 \\ 1 & \operatorname{Im} z < 0. \end{cases}$$

To proceed we need

LEMMA 20.8. Let A be closed, densely defined and symmetric. Then $\operatorname{ran}(A \pm iI)$ are closed.

PROOF. Note that

$$\|(A \pm iI)u\|^2 = \langle Au, Au \rangle \mp i \langle Au, u \rangle \pm i \langle u, Au \rangle \langle u, u \rangle = \|Au\|^2 + \|u\|^2. \quad \square$$

Since $\ker(A^\dagger - i) = \{0\}$ and $\operatorname{ran}(A + iI)$ is closed we conclude that

$$\operatorname{ran}(A + iI) = H.$$

If $B \supset A$ is a self-adjoint extension and $v \in \mathcal{D}(B) \setminus \mathcal{D}(A)$, then there is $u \in \mathcal{D}(A)$ such that

$$Au + iu = Bv + iv.$$

Since B extends A , we conclude that

$$(B + i)(v - u) = 0.$$

This is a contradiction since B is self-adjoint. So no such extension can exist. $\square$

LECTURE 21

Theory of self-adjoint extensions

THEOREM 21.1. *Let B be a closed densely defined symmetric operator. Then for any $z \in \mathbb{C} \setminus \mathbb{R}$, $B - z : \mathcal{D}(B) \rightarrow H$ is injective with closed range and $(B - z)^{-1} : \text{ran}(B - z) \rightarrow H$ is a bounded map with norm less than $1/|\text{Im } z|$.*

PROOF. Let $u \in \mathcal{D}(B)$ and $z \in \mathbb{C}$. Then

$$|\text{Im } z| \|u\|^2 = |\text{Im } \langle (B - z)u, u \rangle| \leq \|(B - z)u\| \|u\|.$$

Thus

$$\|u\| \leq \frac{1}{|\text{Im } z|} \|(B - z)u\|.$$

It follows that $B - z$ is one-to-one and if $f_n \in \text{ran}(B - z)$ is a convergent sequence, then u_n converges where $f_n = (B - z)u_n$. Since B is closed, $u = \lim_n u_n \in \mathcal{D}(B)$ and so $f = \lim_n f_n \in \text{ran } B$. $\square$

Although $B - z$ is injective with closed range, it may still happen that $z \in \sigma(B)$ if $\text{ran}(B - z) \neq H$. Indeed,

THEOREM 21.2. *A closed symmetric operator B is self-adjoint if and only if $\sigma(B) \subset \mathbb{R}$.*

PROOF. The implication $\implies$ was already established. For the converse suppose B is closed symmetric and $\pm i \notin \sigma(B)$. Let $u \in H$ and let $x = (A - i)^{-1}u$. Then

$$\langle u, y \rangle = \langle (A - i)x, y \rangle = \langle x, (A + i)y \rangle = \langle (A - i)^{-1}u, (A + i)y \rangle$$

for all $y \in \mathcal{D}(A)$. Plug in $y = (A + i)^{-1}v$ for some $v \in H$ to learn that

$$\langle u, (A + i)^{-1}v \rangle = \langle (A - i)^{-1}u, v \rangle \quad \text{for all } u, v \in H.$$

That is $[(A + i)^{-1}]^\dagger = (A - i)^{-1}$. Now suppose $v \in \mathcal{D}(A^\dagger)$. Then

$$\langle (A - i)x, v \rangle = \langle x, (A^\dagger + i)v \rangle,$$

for all $x \in \mathcal{D}(A)$. Setting $u = (A - z)x$ we can write this as

$$\langle u, v \rangle = \langle (A - i)^{-1}u, (A^\dagger - i)v \rangle = \langle u, (A + i)^{-1}(A^\dagger + i)v \rangle,$$

for all $u \in H$. Thus

$$v = (A + i)^{-1}(A^\dagger + i)v$$

which shows that $v \in \mathcal{D}(A)$ and that

$$Av + iv = A^\dagger v + iv \implies Av = A^\dagger v. \quad \square$$

THEOREM 21.3. *Let C be a densely defined, closed symmetric operator in a Hilbert space H . Then $\text{codim } \text{ran}(C - z) = \dim \ker(C^\dagger - \bar{z})$ is the same for all z with $\text{Im } z > 0$, and similarly for all z with $\text{Im } z < 0$.*

Let $V_+ = \ker(C^\dagger + i)$, $n_+ = \dim V_+$, and $V_- = \ker(C^\dagger - i)$, $n_- = \dim V_-$. Then the set of self-adjoint extensions to C is in one-to-one correspondence with the set of linear isometries from $U : V_- \rightarrow V_+$. In particular, C has a self-adjoint extension if and only if $n_+ = n_-$.

REMARK. $n_\pm$ are called the deficiency indices of C .

We will use:

LEMMA 21.4. Let V_1, V_2 be closed subspaces of a Hilbert space H . If $\dim V_1 > \dim V_2$ then there is $u \in V_1$ such that $u \neq 0$ and $\langle u, v \rangle = 0$ for all $v \in V_2$.

EXERCISE 53. Prove this.

PROOF. We will prove that $\text{codim } \text{ran}(C - z)$ is constant in the upper and lower half planes next time.

Let V be the Cayley transform of C ,

$$V = (C - i)(C + i)^{-1}, \quad V : \text{ran}(C + i) \rightarrow \text{ran}(C - i).$$

I claim this map is an isometry. Indeed, if $x = (C + i)u$, then $Vx = (C - i)u$, so

$$\begin{aligned} \|Vx\|^2 &= \|(C - i)u\|^2 = \langle Cu, Cu \rangle - i\langle u, Cu \rangle + i\langle Cu, u \rangle + \langle u, u \rangle \\ &= \langle Cu, Cu \rangle + i\langle u, Cu \rangle - i\langle Cu, u \rangle + \langle u, u \rangle = \|(C + i)u\|^2 = \|x\|^2. \end{aligned}$$

We want to extend V to be a unitary map from $H \rightarrow H$. To do this we need an isometry

$$T : V_- = (\text{ran}(C + i))^\perp \rightarrow (\text{ran}(C - i))^\perp = V_+.$$

Let T be such an isometry, consider the unitary map

$$Ux = VPx + T(I - P)x,$$

with P orthogonal projection onto $\text{ran}(C + i)$, and define

$$A = i(I + U)(I - U)^{-1},$$

with domain

$$\mathcal{D}(A) = \text{ran}(I - U).$$

I claim that A is a self-adjoint extension of C .

First we must see that $I - U$ is one-to-one. Suppose on the contrary that

$$x - Ux = 0$$

for some x . Then $x \perp \text{ran}(I - U^\dagger) = \text{ran}(I - U^\dagger)U = \text{ran}(U - I)$. However,

$$\text{ran}(U - I) \supset \text{ran}(V - I),$$

and

$$V - I = (C - i)(C + i)^{-1} - (C + i)(C + i)^{-1} = -2i(C + i)^{-1}.$$

So $\text{ran}(V - I) = \mathcal{D}(C)$ is dense. Thus $x = 0$. This argument also shows that

$$\mathcal{D}(A) = \text{ran}(I - U) \supset \text{ran}(I - V) = \mathcal{D}(C).$$

Let $u \in \mathcal{D}(C)$ and let $x = Cu + iu$. Then

$$(I - U)x = x - Vx = 2i(C + i)^{-1}x = 2iu,$$

$$(I + U)x = x + Vx = 2C(C + i)^{-1}x = 2Cu.$$

Putting these things together we see that A is an extension of C :

$$Au = i(I + U)(I - U)^{-1}u = \frac{1}{2}(I + U)x = C(C + i)^{-1}x = Cu.$$

Thus A is an extension of C .

Self-adjointness of A follows from the spectral representation of U since the function $i(1 + e^{i\theta})(1 - e^{i\theta})^{-1}$ is real valued. However, we can also proceed more directly. Note that

$$\mathcal{D}(A) = \mathcal{D}(C) + (I - T)V_- = \{u + v - Tv : u \in \mathcal{D}(C) \text{ and } v \in V_-\},$$

and since $v - Tv = (I - U)v$, $v \in V_-$, we have

$$A(v - Tv) = i(I + U)v = i(v + Tv).$$

Thus, since $C^\dagger v = iv$ and $C^\dagger Tv = -iTv$,

$$A(u + v - Tv) = Cu + C^\dagger(v - Tv).$$

Let us see that A is symmetric. Indeed,

$$\begin{aligned} \langle A(u + v - Tv), u' + v - Tv \rangle \\ &= \langle Cu, u' \rangle + \langle Cu, v' - Tv' \rangle + \langle C^\dagger(v - Tv), u' \rangle + \langle C^\dagger(v - Tv), (v' - Tv') \rangle \\ &= \langle u, Cu' \rangle + \langle u, C^\dagger(v' - Tv') \rangle + \langle v - Tv, Cu' \rangle + i \langle v + Tv, v' - Tv' \rangle. \end{aligned}$$

The last term satisfies

$$i \langle v + Tv, v' - Tv' \rangle = i \langle Tv, v' \rangle - i \langle v, Tv' \rangle = -\langle Tv, iv' \rangle + \langle v, iTv' \rangle = \langle v - Tv, i(v' + Tv') \rangle$$

since T is an isometry. Symmetry follows.

It remains to show that $\sigma(A) \subset \mathbb{R}$. Suppose $z \in \mathbb{C} \setminus \mathbb{R}$. Then

$$A - zI = i(I + U + iz(I - U))(I - U)^{-1} = i((1 + iz)I + (1 - iz)U)(I - U)^{-1}.$$

Thus

$$(A - zI)^{-1} = -i(I - U)((1 + iz)I + (1 - iz)U)^{-1}.$$

The first factor is clearly bounded and the second factor is bounded since $(1 + iz)/(1 - iz)$ is not in the unit circle and U is unitary. Thus $(A - zI)^{-1}$ is bounded.

EXERCISE 54. Show that any self-adjoint extension A is of the above form. (Hint: show that the Cayley transform of A , when restricted to $\text{ran}(C + i)$, must agree with V .) $\square$

LECTURE 22

Self-adjoint extensions (II) and the Friedrichs extension

1. Proof that $\text{codim ran}(C - z)$ is constant in the upper and lower half planes

Let C be a densely defined, closed symmetric operator on a Hilbert space. Let $V_z = \ker(C^\dagger - \bar{z}) = (\text{ran}(C - z))^\perp$, and define

$$\delta(z; w) = \sup_{v \in V_z; v \neq 0} \inf_{u \in V_w} \frac{\|v - u\|}{\|v\|} = \sup_{v \in V_z: \|v\|=1} \text{dist}(v, V_w).$$

Note that $\delta(z; w) \leq 1$. Furthermore, by Lemma 21.4 from last time, if $\dim V_z > \dim V_w$ then $\delta(z; w) = 1$ since then there is $v \in V_z \cap V_w^\perp$, so that

$$\|v - u\| = \sqrt{\|v\|^2 + \|u\|^2}, \quad u \in V_w.$$

Similarly, let

$$\tilde{\delta}(w; z) = \sup_{v \in V_w^\perp: \|v\|=1} \text{dist}(v, V_z^\perp).$$

Claim: $\tilde{\delta}(w; z) = \delta(z; w)$.

Indeed, for any subspace X

$$\text{dist}(v, X) = \|P_{X^\perp} v\|,$$

with $P_{X^\perp}$ = projection onto $X^\perp$. Thus

$$\tilde{\delta}(w; z) = \sup_{v \in V_w^\perp: \|v\|=1} \sup_{u \in V_z: \|u\|=1} |\langle u, v \rangle| = \sup_{u \in V_z: \|u\|=1} \sup_{v \in V_w^\perp: \|v\|=1} |\langle u, v \rangle| = \sup_{u \in V_z: \|u\|=1} \text{dist}(u, V_w) = \delta(z; w).$$

Now consider the map

$$(C - w)(C - z)^{-1} : \text{ran}(C - z) \rightarrow \text{ran}(C - w).$$

(Recall that $\text{ran}(C - z)$ and $\text{ran}(C - w)$ are closed subspaces.) Let $u \in \text{ran}(C - z)$. Then

$$(C - w)(C - z)^{-1}u = u + (z - w)(C - z)^{-1}u,$$

from which we conclude that

$$\|(C - w)(C - z)^{-1}u - u\| \leq \frac{|z - w|}{|\text{Im } z|} \|u\|.$$

As this holds for all $u \in \text{ran}(C - z) = V_z^\perp$, we conclude that if $|z - w|/|\text{Im } z| < 1$ then

$$\delta(w; z) = \tilde{\delta}(z; w) < 1.$$

Thus $\dim V_w \leq \dim V_z$. Hence, if $|z - w| < \min(|\text{Im } z|, |\text{Im } w|)$ then $\dim V_w = \dim V_z$. It follows that $\dim V_z$ is constant as z ranges through the upper and lower half planes.

REMARK. A more succinct proof $\text{codim ran}(C - z)$ can be made using *index theory*. The basic ideas are as follows:

- (1) $\dim \ker(C - z) = 0$ for all $z \in \mathbb{C} \setminus \mathbb{R}$.

(2) Thus the index of $C - z$,

$$\text{ind}(C - z) = \dim \ker(C - z) - \text{codim} \text{ran}(C - z) = \text{codim} \text{ran}(C - z).$$

This index could be infinity.

(3) There is a general theorem which states that the index of a map is “homotopy invariant” — that is it does not change under continuous perturbations. Here we consider $C - z$ as a linear map from the Hilbert space $\mathcal{D}(C)$, with inner product $\langle u, v \rangle_+ = \langle Cu, Cv \rangle + \langle u, v \rangle$, into H . Thus $z \mapsto C - z$ is a continuous map from $\mathbb{C} \setminus \mathbb{R}$ into the bounded operators from $\mathcal{D}(C) \rightarrow H$. (Essentially, we proved a special case of that theorem.)

This is all quite elementary, but we haven’t covered the relevant index theory.

2. Friedrichs extension

DEFINITION 22.1. A symmetric operator A defined on a dense subspace $\mathcal{D}$ is semibounded if

$$c \|u\|^2 \leq \langle Au, u \rangle$$

for some constant c and for all $u \in \mathcal{D}$.

THEOREM 22.1 (Friedrichs). *Let A be a densely defined symmetric operator, semi-bounded with constant c . Then A has a self-adjoint extension with $\sigma(A) \subset [c, \infty)$.*

PROOF. We could evaluate deficiency indices, but the point here is actually the construction. So let us proceed directly. First of all we will assume that $c = 1$. This can be achieved without loss of generality by replacing A by $(1 - c)I + A$.

The expression

$$\langle u, v \rangle_+ = \langle Au, v \rangle$$

defines an inner product on $\mathcal{D}(A)$. Furthermore

$$\|u\|_+^2 = \langle Au, u \rangle \geq \|u\|^2$$

by assumption. Thus convergence in $\|\cdot\|_+$ controls convergence in H . Let $\mathcal{Q}(A)$ denote the Hilbert space completion of $\mathcal{D}(A)$ in this norm. By the norm domination, we can take this to be a subspace of H : $\mathcal{D}(A) \subset \mathcal{Q}(A) \subset H$.

Now let $x \in H$. Then

$$y \mapsto \langle y, x \rangle$$

is a bounded linear functional on $\mathcal{Q}(A)$. Thus there is a unique element $x' \in \mathcal{Q}(A)$ such that

$$\langle y, x \rangle = \langle y, x' \rangle_+ \quad \text{for all } y \in \mathcal{Q}(A).$$

Let $T : H \rightarrow \mathcal{Q}(A)$ denote the map $Tx = x'$. Furthermore, considered as a map $T : H \rightarrow H$, T is bounded. Indeed,

$$\|Tx\| \leq \|Tx\|_+ = \sup_{y \in \mathcal{Q}(A)} \frac{|\langle y, Tx \rangle_+|}{\|y\|_+} \leq \sup_y \frac{|\langle y, x \rangle|}{\|y\|} = \|x\|$$

and we see that $\|T\| \leq 1$.

Let $\mathcal{D} = \text{ran } T$ and define $B : \mathcal{D} \rightarrow H$ to be the inverse map T^{-1} . This is a good definition since if $Tx = 0$ then $\langle y, x \rangle = 0$ for all $y \in \mathcal{Q}(A)$ which is a dense subset of H (since $\mathcal{D}(A)$ is) and thus $x = 0$. For $v \in \mathcal{D}$ we have

$$\langle u, v \rangle_+ = \langle u, Bv \rangle \quad \text{for all } u \in \mathcal{Q}(A).$$

Thus

$$\langle u, Bv \rangle = \langle u, v \rangle_+ = \overline{\langle v, u \rangle_+} = \overline{\langle v, Bu \rangle} = \langle Bu, v \rangle$$

for $u, v \in \mathcal{D}$. That is B is symmetric. It follows that

$$\langle x, Ty \rangle = \langle BTx, Ty \rangle = \langle Tx, BTy \rangle = \langle Tx, y \rangle$$

for all $x, y \in H$. Thus T is symmetric. Since T is also bounded, it is self-adjoint.

Since $\|T\| \leq 1$ the spectrum of T is contained in $[-1, 1]$. However, since

$$\langle Tx, x \rangle = \langle Tx, BTx \rangle = \|Tx\|_+^2 \geq 0$$

we see that $\sigma(T) \subset [0, 1]$. Thus for any $z \in \mathbb{C} \setminus [1, \infty)$, $(I - zT)$ has a bounded inverse, and furthermore,

$$(I - zT)^{-1}T(B - z) = I = (B - z)(I - zT)^{-1}T.$$

Thus $\sigma(B) \subset [1, \infty) \subset \mathbb{R}$ and so B is self-adjoint.

Now if $x \in \text{ran}(A)$, say $x = Ax'$ then

$$\langle y, x \rangle = \langle y, Ax' \rangle = \langle y, x' \rangle_+ \quad \text{for all } y \in \mathcal{D}(A).$$

Since $\mathcal{D}(A)$ is dense in $\mathcal{Q}(A)$ we find that

$$\langle y, x \rangle = \langle y, A^{-1}x \rangle_+$$

for all $x \in \text{ran } A$ and $y \in \mathcal{Q}(A)$. (Note that A is one-to-one by the assumed lower bound.)

Thus $\mathcal{D}(A) \subset \mathcal{D} = \text{ran } T$ and

$$Bx' = Ax' \quad \text{for all } x' \in \mathcal{D}(A).$$

So B extends A . □

The operator B constructed in the proof is known as the Friedrichs extension of A .

EXERCISE 55. Let $A = -\frac{d^2}{dx^2}$ on $C_c^2(0, 1) \subset L^2(0, 1)$. Show that A is densely defined, symmetric and non-negative ($\langle Au, u \rangle \geq 0$). What is the Friedrich's extension of A ? Does A have other self-adjoint extensions? If yes, what are they?

Part 3

Semigroups

LECTURE 23

Continuous and strongly continuous semigroup

DEFINITION 23.1. A one parameter semigroup of operators is a map $Z : [0, \infty) \rightarrow \mathcal{L}(X)$, with X a Banach space, and such that

$$Z(t+s) = Z(t)Z(s) \quad \text{for all } t, s \geq 0,$$

and with $Z(0) = I$.

THEOREM 23.1. *Let $Z(t) : X \rightarrow X$ be a one-parameter semigroup of operators that is norm continuous of zero:*

$$\lim_{t \rightarrow 0} \|Z(t) - I\| = 0.$$

Then

- (1) $Z(t)$ is a norm continuous for all $t \geq 0$.
- (2) There is a bounded linear map $G \in \mathcal{L}(X)$ such that

$$Z(t) = e^{tG},$$

where the exponential is defined by the power series

$$e^{tG} = \sum_{n=0}^{\infty} \frac{1}{n!} t^n G^n.$$

Conversely, given any bounded linear map $G \in \mathcal{L}(X)$, e^{tG} defines a norm continuous one-parameter semigroup.

REMARK. One could also define e^{tG} by the Riesz integral:

$$e^{tG} = \frac{1}{2\pi i} \oint_{\gamma} \frac{1}{zI - G} e^{tz} dz.$$

PROOF. The converse statement is a consequence of the Riesz functional calculus for operators in a Banach algebra.

Suppose $Z(t)$ is a one parameter semigroup continuous at 0. Since

$$Z(t+h) - Z(t) = (Z(h) - I)Z(t),$$

we see that $Z(t)$ is continuous at every $t \geq 0$.

There is $\epsilon > 0$ such that for $t < \epsilon$

$$\|Z(t) - I\| < 1.$$

We will use

LEMMA 23.2. *Let Z be an operator with $\|Z - I\| < 1$. Then*

$$\log Z = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (Z - I)^n$$

is a uniformly convergent series and

$$e^{\log Z} = Z.$$

Furthermore, if Z, W are operators such that

- (1) $\|ZW - I\| < 1$, $\|Z - I\| < 1$, $\|W - I\| < 1$ and
- (2) $ZW = WZ$

then

$$\log(ZW) = \log Z + \log W.$$

EXERCISE 56. Prove the lemma

It follows that

$$L(t) = \log(Z(t)) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} [Z(t) - I]^n$$

is a uniformly convergent series $t < \epsilon$,

$$Z(t) = e^{L(t)}$$

and

$$L(t+s) = L(t) + L(s)$$

if $t+s < \epsilon$. We conclude that for rational $t < \epsilon$,

$$\frac{1}{t}L(t)$$

is independent of t . Let the operator be G . So

$$L(t) = tG, \quad t \in \mathbb{Q} \cap [0, \epsilon).$$

By continuity

$$L(t) = tG \quad \forall t \in [0, \epsilon).$$

Thus

$$Z(t) = e^{tG} \quad \forall t \in [0, \epsilon).$$

For arbitrary $t = n\epsilon + r$ the result follows from the multiplicative property. $\square$

DEFINITION 23.2. A one parameter semigroup is strongly continuous at $t = 0$ if

$$\lim_{t \rightarrow 0} \|Z(t)x - x\| = 0$$

for all $x \in X$.

THEOREM 23.3. Let $Z(t)$ be a one-parameter semigroup of operators that is strongly continuous at $t = 0$.

- (1) There exist constants b and k such that $Z(t)$ is bounded in norm by

$$\|Z(t)\| \leq be^{kt}.$$

- (2) $Z(t)x$ is a continuous function of t for each $x \in X$.

PROOF. Part 2 — continuity of $Z(t)x$ for all t follows — since

$$Z(t+h)x - Z(s)x = Z(s)[Z(t-s)x - x].$$

To prove (1) we will show first that $Z(t)$ is uniformly bounded for t in a neighborhood of 0. Suppose on the contrary that there is a sequence $t_n \rightarrow 0$ such that

$$\|Z(t_n)\| \rightarrow \infty.$$

By the principle of uniform boundedness we conclude that

$$\|Z(t_n)x\| \rightarrow \infty$$

for some $x \in X$. But $Z(t_n)x \rightarrow x$ so this is a contradiction. Thus given $\epsilon > 0$,

$$b := \sup_{0 \leq t \leq \epsilon} \|Z(t)\| < \infty.$$

Clearly $b \geq 1$. Now for arbitrary $t = n\epsilon + r$, we have $Z(t) = Z(r)[Z(\epsilon)]^n$, so

$$\|Z(t)\| \leq b^{n+1} = be^{n \ln b} \leq be^{kt},$$

with $k = \epsilon^{-1} \ln b$. □

LECTURE 24

Unbounded operators in a Banach space

Our main goal in the coming lectures will be to prove a formula that is of the form

$$Z(t) = e^{tG}$$

for a strongly continuous semi-group. Clearly the operator G cannot be bounded (unless the semi-group is in fact norm continuous). Thus we must consider unbounded operators in a Banach space setting. So our first task is to understand some generalizations of the notion of unbounded operators to a Banach space. The first idea — and ultimately the right one — is to define

$$Gx = \partial_t Z(t)x \big|_{t=0}.$$

However, this cannot work for every x since

THEOREM 24.1. *Let $Z(t)$ be a strongly continuous semigroup on a Banach space such that*

$$Gx = \partial_t Z(t)x \big|_{t=0}$$

exists for every $x \in X$. Then G is bounded and $Z(t) = e^{tG}$ is norm continuous.

REMARK. This is analogous to the Hellinger-Toeplitz theorem. It shows that the generator of a semigroup that is strongly continuous but not norm continuous cannot be everywhere defined.

PROOF. Consider the family of operators

$$D_h = \frac{1}{h}(Z(h) - I).$$

Since $\lim_{h \rightarrow 0} D_h x = Gx$ exists for every x we see that

$$\sup_{h \in [0,1]} \|D_h x\| \leq c(x).$$

By the principle of uniform boundedness, there is a constant C such that

$$\sup_{h \in [0,1]} \|D_h\| \leq C.$$

Thus $\|G\| \leq C$. □

The following definitions are analogous to the definitions in the Hilbert space context:

DEFINITION 24.1. A densely defined linear operator G on a Banach space X is a linear map G defined on a dense subspace $\mathcal{D}(G) \subset X$ called the domain of G . We say that G is closed if the graph $\{(x, Gx) : x \in \mathcal{D}(G)\}$ is closed, in other words if whenever $x_n \rightarrow x$ and $Gx_n \rightarrow y$ then $y \in \mathcal{D}(G)$ and $Gx = y$.

Let $\overline{\mathcal{L}}(X)$ denote the set of densely defined closed linear maps of X into itself.

DEFINITION 24.2. Let $G \in \overline{\mathcal{L}}(X)$. The resolvent set $\rho(G)$ is the collection of complex numbers ζ such that $\zeta I - G$ maps $\mathcal{D}(G)$ one-to-one onto X . The spectrum of G , $\sigma(G)$, is the complement of the resolvent set.

If $\zeta \in \rho(G)$ then $\zeta I - G$ is invertible, it's inverse

$$R(\zeta) = (\zeta I - G)^{-1}$$

is an everywhere defined closed linear map and is thus bounded by the closed graph theorem.

DEFINITION 24.3. Let $Z(t)$ be a strongly continuous one parameter semi-group on a Banach space X . The infinitesimal generator of $Z(t)$ is the linear map

$$Gu = \lim_{t \rightarrow 0} \frac{1}{t} [Z(t)u - u],$$

defined on the domain $\mathcal{D}(G)$ consisting of vectors u such that the right hand side exists (in norm).

THEOREM 24.2. Let G be the infinitesimal generator of a strongly continuous semi-group. Then $G \in \overline{\mathcal{L}}(X)$. Furthermore

- (1) $Z(t) : \mathcal{D}(G) \rightarrow \mathcal{D}(G)$ and G commutes with $Z(t)$.
- (2) For each $n \geq 0$, G^n is densely defined.
- (3) If $\|Z(t)\| \leq be^{kt}$ then all complex numbers ζ with $\operatorname{Re} \zeta > k$ belong to $\rho(G)$.
- (4) The resolvent of G is the Laplace transform of $Z(t)$,

$$R(\zeta) := (\zeta I - G)^{-1} = \int_0^\infty e^{-\zeta t} Z(t) dt.$$

PROOF. Recall that we showed that $Z(t)$ is strongly continuous and bounded by $\|Z(t)\| \leq be^{kt}$ for suitable b, k . Thus

$$I(t)x = \int_0^t Z(s)x ds$$

defines a bounded operator, with the integral on the l.h.s. a Riemann integral.

Claim: $I(t) : X \rightarrow \mathcal{D}(G)$.

Indeed,

$$\frac{1}{h} [Z(h) - I] I(t)x = \frac{1}{h} \int_0^t [Z(s+h)x - Z(s)x] ds = \frac{1}{h} \int_t^{t+h} Z(s)x ds - \frac{1}{h} \int_0^h Z(s)x ds.$$

It follows that $I(t)x \in \mathcal{D}(G)$ and

$$GI(t)x = Z(t)x - x.$$

In particular, since $\lim_{t \rightarrow 0} I(t)x = x$ we see that $\mathcal{D}(G)$ is dense.

Now since

$$\frac{1}{h} [Z(t+h) - Z(t)]x = Z(t) \frac{1}{h} [Z(h) - I]x = \frac{1}{h} [Z(h) - I] Z(t)x$$

we see that $Z(t)$ maps $\mathcal{D}(G) \rightarrow \mathcal{D}(G)$ and

$$\frac{d}{dt} Z(t)x = Z(t)Gx = GZ(t)x.$$

for such x . (1) follows as does the identity

$$Z(t)x - x = \int_0^t Z(s)Gx \, ds$$

for $x \in \mathcal{D}(G)$.

Now suppose $x_n \in \mathcal{D}(G)$ and $x_n \rightarrow x$ and $Gx_n \rightarrow y$. Then

$$Z(t)x_n - x_n = \int_0^t Z(s)Gx_n \, ds.$$

Taking limits on both sides we conclude that

$$Z(t)x - x = \int_0^t Z(s)y \, ds.$$

Dividing by t and taking the limit $t \rightarrow 0$ we see that $x \in \mathcal{D}(G)$ and $Gx = y$. That is G is closed.

Similarly, if $\phi : (0, \infty) \rightarrow \mathbb{R}$ is infinitely differentiable and compactly supported then we may define

$$I_\phi x = \int_0^\infty \phi(s)Z(s)x \, ds.$$

Arguing as above we conclude that $I_\phi : X \rightarrow \mathcal{D}(G)$ and

$$GI_\phi x = - \int_0^\infty \phi'(s)Z(s)x \, ds = -I_{\phi'} x.$$

Thus $\text{ran } I_\phi \subset \mathcal{D}(G^n)$. Taking $\phi_j \rightarrow \delta$ with $\phi_j \in C^\infty$ we see that $\mathcal{D}(G^n)$ is dense.

Now suppose $\text{Re } \zeta > k$ with k as in part (3). Then

$$L(\zeta)x = \int_0^\infty e^{-\zeta t} Z(t)x \, dt$$

is a convergent (improper) integral. Since

$$\begin{aligned} \frac{1}{h} [Z(h) - I] L(\zeta)x &= \frac{1}{h} \int_0^\infty e^{-\zeta t} [Z(t+h) - Z(t)] x \, dt \\ &= -\frac{1}{h} \int_0^h Z(t)x \, dt + \frac{1}{h} \int_h^\infty [e^{-\zeta(t-h)} - e^{-\zeta t}] Z(t)x \, dt. \\ &\rightarrow -x + \int_0^\infty \zeta e^{-\zeta t} Z(t)x \, dt. \end{aligned}$$

It follows that $L(\zeta)x \in \mathcal{D}(G)$ and

$$GL(\zeta)x = -x + \zeta L(\zeta)x.$$

Likewise,

$$L(\zeta)Gx = -x + \zeta L(\zeta)x.$$

Thus $\zeta \in \rho(G)$ and

$$R(\zeta) = L(\zeta). \quad \square$$

LECTURE 25

The Hille-Yosida theorem

THEOREM 25.1. *Let $Z(t)$ and $\tilde{Z}(t)$ be strongly continuous semigroups with generators G and $\tilde{G}$. If $G = \tilde{G}$ then $Z(t) = \tilde{Z}(t)$ for all $t \geq 0$.*

PROOF. Let $G = \tilde{G}$ and let $x \in \mathcal{D}(G)$. Then

$$\partial_t W(t)Z(s-t)x = W(t)GZ(s-t)x - W(t)GZ(s-t)x = 0.$$

It follows that

$$W(s)x - Z(s)x = \int_0^s \partial_t W(t)Z(s-t)x = 0.$$

Since $W(s)$ and $Z(s)$ are bounded and agree on the dense set $\mathcal{D}(G)$, they are equal. $\square$

We now turn our attention to semigroups which satisfy

$$\|Z(t)\| \leq e^{-kt}$$

with some $k \in \mathbb{R}$. Since $e^{kt}Z(t)$ is a semi-group, it suffices to consider

$$\|Z(t)\| \leq 1,$$

that is $Z(t)$ is a semi-group of contractions.

THEOREM 25.2 (Hille-Yosida). *Let $Z(t)$ be a strongly continuous semigroup of contractions with generator G . Then*

- (1) $(0, \infty) \subset \rho(G)$, and
- (2) for any $\lambda \in (0, \infty)$

$$\|(\lambda I - G)^{-1}\| \leq \frac{1}{\lambda}.$$

Conversely, if $G \in \overline{\mathcal{L}}(X)$ satisfies (1), (2) then it is the generator of a strongly continuous semigroup.

PROOF. We proved (1) and (2) last time under the more general condition that $\|Z(t)\| \leq be^{-kt}$. Recall

$$(\lambda I - G)^{-1} = \int_0^\infty e^{-\lambda t} Z(t) dt.$$

Turning to the converse, let $R(n) = (nI - G)^{-1}$. We begin by showing that

$$\lim_n nR(n)x = x$$

for all $x \in X$. To see this, note that

$$nR(n) - I = R(n)G.$$

Thus for $x \in \mathcal{D}(G)$

$$\|nR(n)x - x\| \leq \frac{1}{n} \|Gx\|.$$

Thus $nR(n)x \rightarrow x$ for all $x \in \mathcal{D}(G)$. By an $\epsilon/2$ argument it follows that $nR(n)x \rightarrow x$ for all $x \in X$.

Now let

$$G_n = nGR(n).$$

Since

$$G_n = n^2R(n) - nI,$$

we see that G_n is bounded

$$\|G_n\| \leq 2n.$$

Note that

$$Gx = \lim_n G_n x \quad x \in \mathcal{D}(G).$$

Consider the semigroups $Z_n(t) = e^{tG_n}$,

$$Z_n(t) = e^{-nt} e^{tn^2R(n)} = e^{-nt} \sum_{m=0}^{\infty} \frac{(n^2t)^m}{m!} [R(n)]^m.$$

Thus

$$\|Z_n(t)\| \leq e^{-nt} \sum_{m=0}^{\infty} \frac{(n^2t)^m}{m!} \frac{1}{n^m} = e^{-nt} e^{nt} = 1.$$

Fix n, m integers and let $x \in \mathcal{D}(G)$. Consider

$$\partial_t Z_n(s-t)Z_m(t)x = Z_n(s-t)(G_n - G_m)Z_m(t)x = Z_n(s-t)Z_m(t)(G_n - G_m)x,$$

where in the last step we have noted that G_n and G_m commute with $Z_m(t)$. Integrating over t from 0 to s we see that

$$Z_m(s)x - Z_n(s)x = \int_0^s Z_n(s-t)Z_m(t)(G_n - G)x ds.$$

Thus

$$\|Z_m(s)x - Z_n(s)x\| \leq s \|G_n x - G_m x\|.$$

Since $G_n x \rightarrow Gx$ we see that $Z_n(s)x$ is a Cauchy sequence. Denote its limit by $Z(s)x$. The operator $Z(s)$ extends to all of X by an approximation argument since

$$\|Z(s)x\| \leq \|x\| \quad x \in \mathcal{D}(G).$$

As the Cauchy estimate is locally uniform in s , we get uniform convergence of $Z_n(s)x \rightarrow Z(s)x$. Thus $Z(s)$ is strongly continuous. Clearly $Z(s) = I$. Finally

$$Z(t+s)x = \lim_n Z_n(t)Z_n(s)x = Z(t)Z(s)x$$

so Z is a semi-group. Clearly $\|Z(t)\| \leq 1$.

It remains to show that G is the generator of $Z(t)$. Since

$$Z_n(t)x - x = \int_0^t Z_n(s)G_n x ds.$$

Suppose $x \in \mathcal{D}(G)$. Then

$$Z(t)x - x = \int_0^t Z(s)Gx ds.$$

It follows that the generator H of $Z(t)$ extends G , that is $\mathcal{D}(G) \subset \mathcal{D}(H)$ and $Hx = Gx$ for all $x \in \mathcal{D}(G)$. By the theorem of last lecture, all ζ with $\operatorname{Re} \zeta > 0$ are in the resolvent set of H . Since $\zeta I - G$ is surjective and $\zeta I - H$ is one-to-one it follows that $\mathcal{D}(G) = \mathcal{D}(H)$. $\square$

THEOREM 25.3 (Lumer-Phillips). *Let G be a densely defined operator on a Hilbert space H whose resolvent set includes $[0, \infty)$. Then G is the generator of a semigroup of contractions if and only if G is “dissipative:”*

$$\operatorname{Re} \langle Gx, x \rangle \leq 0$$

for all $x \in \mathcal{D}(G)$.

PROOF. Suppose $\operatorname{Re} \langle Gx, x \rangle \leq 0$ for all $x \in \mathcal{D}(G)$. Since

$$\|\lambda x - Gx\|^2 = \lambda^2 \|x\|^2 + \|Gx\|^2 - 2 \operatorname{Re} \langle Gx, x \rangle,$$

we conclude that

$$\|x\|^2 \leq \frac{1}{\lambda^2} \|\lambda x - Gx\|^2$$

for all $x \in \mathcal{D}(G)$. It follows that

$$\|R(\lambda)x\| \leq \frac{1}{\lambda} \|x\| \quad \text{for all } x,$$

which is $\|R(\lambda)\| \leq 1/\lambda$.

Conversely, if $\|R(\lambda)\| \leq 1/\lambda$ then

$$\|x\|^2 \leq \frac{1}{\lambda^2} \|\lambda x - Gx\|^2$$

for all $x \in \mathcal{D}(G)$. Cancelling $\|x\|^2$ from both sides, we conclude that

$$2 \operatorname{Re} \langle Gx, x \rangle \leq \frac{1}{\lambda^2} \|Gx\|^2. \quad \square$$

LECTURE 26

Unitary semigroups and self-adjoint semigroups

A particular example of a contractive semigroup is a unitary semigroup $U(t)$. In fact, since a unitary operator is invertible, we can extend a unitary semigroup to be a one-parameter Unitary group by setting $U(-t) = U(t)^{-1} = U(t)^\dagger$ for $t < 0$.

THEOREM 26.1 (Stone and von-Neumann). *Let $U(t)$ be strongly continuous unitary group, with generator*

$$G = \lim_{h \rightarrow 0} [U(h)x - x], \quad (\star)$$

on the domain such that the limit on the r.h.s. exists. Then $-iG$ is self adjoint. Conversely, if A is a self-adjoint operator then iA is the generator of a strongly continuous unitary group.

PROOF. Let $U(t)$ be a unitary group. Then $Z(t) = U(t)$, $t > 0$, is a contractive semigroup, so it has a generator G defined by $(\star)$ with a one-sided limit on the r.h.s.. Since $Z(t)^\dagger = U(-t)$ is also a contractive semigroup, it has a generator which we easily see to be $-G$:

$$\lim_{h \rightarrow 0} \frac{1}{h} [U(-h)x - I] = \lim_{h \rightarrow 0} U(-h) [x - U(h)x] = -Gx$$

for $x \in \mathcal{D}(G)$. By the Hille-Phillips theorem $(0, \infty) \subset \rho(G)$. Likewise $(0, \infty) \subset \rho(-G)$. Let $A = -iG$. Thus $i\mathbb{R} \setminus \{0\} \subset \rho(A)$. Since A is closed and densely defined, to see that it is self-adjoint we need only show that it is symmetric. This, however follows since

$$\begin{aligned} \langle Ax, y \rangle &= -i \langle Gx, y \rangle = -i \lim_{h \rightarrow 0} \frac{1}{h} [\langle U(h)x, y \rangle - \langle x, y \rangle] \\ &= -i \lim_{h \rightarrow 0} \frac{1}{h} [\langle x, U(-h)y \rangle - \langle x, y \rangle] = i \langle x, Gy \rangle = \langle x, Ay \rangle. \end{aligned} \quad (26.1)$$

Now let A be a self-adjoint operator. Then $\mathbb{C} \setminus \mathbb{R} \subset \rho(A)$ and

$$\|(\zeta I - A)^{-1}\| \leq \frac{1}{|\operatorname{Im} \zeta|}.$$

It follows from the Hille-Phillips theorem that iA and $-iA$ generate contractive semigroups. Let them $U(t)$ and $V(t)$ respectively. It follows that

$$\frac{d}{dt} V(t)U(t)x = iV(t)AU(t)x - iV(t)AU(t)x = 0,$$

for $x \in \mathcal{D}(A)$. Since $V(0)U(0)x = x$ we see that $V(t)U(t)x = x$ for all $t \geq 0$. Similarly $U(t)V(t)x = x$ for all $t \geq 0$. Since $U(t)$ and $V(t)$ are bounded operators it follows that $U(t)V(t)x = V(t)U(t)x = x$ for all $x \in H$. That is $V(t) = U(t)^{-1}$. Thus $U(t)$ can be extended to one-parameter group by setting

$$U(t) = V(-t) \quad \text{for } t < 0.$$

Clearly,

$$iAx = \lim_{h \rightarrow 0} \frac{1}{h} [U(h)x - x]$$

for all $x \in \mathcal{D}(A)$. It remains to show that $U(t)$ is unitary. For this purpose, we compute

$$\frac{d}{dt} \|U(t)x\|^2 = \langle -iAU(t)x, U(t)x \rangle + \langle U(t), -iAU(t)x \rangle = 0$$

for $x \in \mathcal{D}(A)$. It follows that $\|U(t)x\| = \|x\|$ for all $x \in H$. $\square$

THEOREM 26.2. *$G \in \overline{\mathcal{L}}(H)$, with H a Hilbert space, is the generator of a strongly continuous semi-group of self-adjoint operators if and only if it is self-adjoint and bounded from above.*

PROOF. Suppose $Z(t)$ is a strongly continuous semigroup of self-adjoint operators with generator G . Let $x, y \in \mathcal{D}(G)$. Then

$$\langle Gx, y \rangle = \lim_{t \downarrow 0} t^{-1} [\langle Z(t)x, y \rangle - \langle x, y \rangle] = \lim_{t \downarrow 0} t^{-1} [\langle x, Z(t)y \rangle - \langle x, y \rangle] = \langle x, Gy \rangle.$$

Thus G is symmetric. We have seen that there are b, k such that $\|Z(t)\| \leq be^{kt}$ and that any ζ with $\operatorname{Re} \zeta > k$ is in $\rho(G)$, since

$$(\zeta I - G)^{-1} = \int_0^\infty e^{-\zeta t} Z(t) dt.$$

Thus the deficiency indices of G are both zero so G is self-adjoint. It follows from the spectral representation of G that

$$\langle Gx, x \rangle \leq k \|x\|^2,$$

so G is bounded from above.

Now suppose G is self-adjoint and bounded from above. Then e^{tG} is well defined by the continuous functional calculus

$$e^{tG}x = \int_{-\infty}^k e^{t\lambda} dE(\lambda)x,$$

where k is an upper bound for G . It follows from the functional calculus that this defines a self-adjoint semigroup which is easily seen to be strongly continuous.

EXERCISE 57. Show that G is the generator of e^{tG} . $\square$

Part 4

Perturbation Theory

LECTURE 27

Rellich's Theorem

THEOREM 27.1. *Let A be a self-adjoint operator in a Hilbert space H with domain $\mathcal{D}(A)$. If T is a symmetric operator on H ,*

- (1) $\mathcal{D}(T) \supset \mathcal{D}(A)$, and
- (2) there are $b < 1$ and $a > 0$ such that

$$\|Tu\| \leq a\|u\| + b\|Au\|,$$

then $A + T$ is self-adjoint on the domain $\mathcal{D}(A)$.

REMARK. If T is bounded then we may take $b = 0$ so $A + T$ is self-adjoint.

PROOF. It is clear that $A + T$ is densely defined and symmetric. Let us show that it is closed. To this end note that by the triangle inequality

$$\|Au\| \leq \|(A + T)u\| + \|Tu\| \leq \|(A + T)u\| + a\|u\| + b\|Au\|$$

for all $u \in \mathcal{D}(A)$. Thus,

$$\|Au\| \leq \frac{1}{1-b} \|(A + T)u\| + \frac{a}{1-b} \|u\|, \quad u \in \mathcal{D}(A).$$

If $u_n \in \mathcal{D}(A)$ converges to u and $(A + T)u_n$ converges to v we see that Au_n is a Cauchy sequence. Since A is closed, we conclude that $u \in \mathcal{D}(A)$ and $Au = \lim_n Au_n$. Since

$$\|Tu - Tu_n\| \leq a\|u - u_n\| + b\|A(u - u_n)\| \rightarrow 0$$

we see that $Tu_n \rightarrow Tu$ and thus that $(A + T)u = v$ so $A + T$ is closed.

To complete the proof of self-adjointness it suffices to show that $\pm ic \in \rho(T + A)$ for some $c > 0$. Let $c \in \mathbb{R} \setminus \{0\}$. Since $A + T$ is closed and symmetric $A + T + ic$ maps $\mathcal{D}(A)$ one-to-one and onto a closed subspace of H . Suppose $v \perp \text{ran}(A + T + ic)$:

$$\langle (A + T + ic)u, v \rangle = 0 \quad \text{for all } u \in \mathcal{D}(A).$$

Since A is self-adjoint, $\text{ran}(A + ic) = H$ and there is $w \in \mathcal{D}(A)$ with $Aw + icw = v$. Thus

$$\langle Aw + icw, Aw + icw \rangle + \langle Tw, Aw + icw \rangle = 0,$$

so

$$\|(A + ic)w\|^2 \leq \|Tw\| \|Aw + icw\|,$$

and

$$\|(A + ic)w\| \leq \|Tw\| \leq a\|w\| + b\|Aw\|.$$

Since A is self-adjoint we find that

$$\|Aw\|^2 + c^2 \|w\|^2 \leq a^2 \|w\|^2 + b^2 \|Aw\|^2 + 2ab \|w\| \|Aw\| \leq \left(a^2 + \frac{ab}{\epsilon}\right) \|w\|^2 + (b^2 + ab\epsilon) \|Aw\|^2.$$

$$\|Aw\|^2 + c^2 \|w\|^2 + \langle Tw, Aw \rangle - ic \langle Tw, w \rangle = 0.$$

Choose ϵ small enough that $b^2 + ab\epsilon < 1$ and choose $c^2 > a^2 + \frac{ab}{\epsilon}$. Then $w = 0$ so $v = 0$ and $\text{ran}(A + T + ic) = H$. Since there are positive and negative choices for c the deficiency indices of $A + T$ are both 0 and it is self-adjoint. $\square$

PROPOSITION 27.2. *Let $n = 1, 2, 3$ or 4 and let $A = -\Delta$ on $L^2(\mathbb{R}^n)$. Suppose $Tf(x) = V(x)f(x)$ with $V \in L^2(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ and real valued. Then $A + T = -\Delta + V$ is self-adjoint on the domain $\mathcal{D}(A)$ of $-\Delta$.*

PROOF.

LEMMA 27.3. *For any n , $-\Delta$ is self-adjoint on the domain $H^2 = \{f \in L^2(\mathbb{R}^n) : \partial_i \partial_j f \in L^2, i, j = 1, \dots, 3\}$.*

PROOF. It is a standard exercise to show that

$$\widehat{\partial_i \partial_j f}(\mathbf{k}) = -4\pi^2 k_i k_j \widehat{f}(\mathbf{k}),$$

where $\widehat{f}$ denotes the Fourier transform of f . In particular

$$\widehat{-\Delta f}(\mathbf{k}) = 4\pi^2 |\mathbf{k}|^2 \widehat{f}(\mathbf{k}).$$

Since Fourier transform is an isometry $L^2 \rightarrow L^2$, we conclude that $-\Delta$ is self-adjoint on the domain

$$H^2 = \{f : |\mathbf{k}|^2 \widehat{f}(\mathbf{k}) \in L^2\},$$

and that this domain is equal to the set given in the Lemma. $\square$

LEMMA 27.4. *Let $V \in L^2 + L^\infty$. Given $\epsilon > 0$ there are $V_2 \in L^2$ and $V_\infty \in L^\infty$ such that $V = V_2 + V_\infty$ and*

$$\|V_2\|_2 < \epsilon.$$

EXERCISE 58. Prove this lemma.

Now given $V \in L^2 + L^\infty$, let $V_{2,\infty}$ be as in the Lemma with ϵ to be chosen below. Then

$$\|Vf\|_2 \leq \|V_\infty f\|_2 + \|V_2 f\|_2 \leq \|V_\infty\|_\infty \|f\|_2 + \|V_2\|_2 \|f\|_\infty$$

for any $f \in L^2 \cap L^\infty$. However, for $n = 1, 2, 3$, or 4 the Sobolev inequalities show that

$$\|f\|_\infty \leq c_n \|-\Delta f\|_2.$$

In particular $\mathcal{D}(-\Delta) \subset L^\infty \cap L^2$ and

$$\|Vf\| \leq a \|f\|_2 + b \|-\Delta f\|_2,$$

with $a = \|V_\infty\|_\infty$ and $b = \epsilon c_n$, with ϵ chosen so that $b < 1$. Thus the hypotheses of Rellich's theorem are satisfied and $T + V$ is self-adjoint. $\square$

EXAMPLE. The operator

$$-\Delta + \frac{1}{|x|}$$

on $L^2(\mathbb{R}^3)$ is the "Hamiltonian for the Hydrogen atom," in suitable units. This is a fundamental operator in quantum mechanics and the above proposition shows it to be self-adjoint since $1/|x|$ is locally L^2 and bounded at ∞ .

PROPOSITION 27.5. *Let $A = -\Delta$ on $L^2(\mathbb{R}^n)$ and let $T = \mathbf{a}(x) \cdot \nabla - \nabla \cdot \mathbf{a}(x)$ with $\mathbf{a}(x)$ a real valued vector field on $\mathbb{R}^n$ with components in $L^\infty(\mathbb{R}^n)$ and such that $\nabla \cdot \mathbf{a} \in L^\infty$. Then $A + T$ is self adjoint on $H^2 = \mathcal{D}(A)$.*

PROOF.

$$\|Tf\|_2 \leq 2 \|\mathbf{a} \cdot \nabla f\|_2 + \|\nabla \cdot \mathbf{a}\| \|f\|_2 \leq 2 \|\mathbf{a}\|_\infty \|\nabla f\|_2 + \|\nabla \cdot \mathbf{a}\|_\infty \|f\|_2.$$

Now

$$\|\nabla f\|_2^2 = \langle -\Delta f, f \rangle \leq \|-\Delta f\| \|f\| \leq \epsilon \|-\Delta f\|^2 + \frac{1}{\epsilon} \|f\|^2.$$

Thus

$$\|Tf\|_2 \leq b \|-\Delta f\|_2 + a \|f\|_2,$$

where we may choose b as small as we like. □

When the estimate

$$\|Tx\| \leq a \|x\| + b \|Ax\|$$

holds, we say that T is A -relatively bounded. The A -relative bound of T is the infimum over b such that the inequality holds for some a . In the last example T has relative bound 0. Thus Rellich's theorem says that if T is symmetric with A -relative bound < 1 and A is self-adjoint then $T + A$ is self-adjoint.

The notion of relative bound holds for densely defined closed operators in a Banach space, but is not as useful. We do have, however,

THEOREM 27.6. *Let $A \in \overline{\mathcal{L}}(X)$ and let T be densely defined with $\mathcal{D}(T) \supset \mathcal{D}(A)$. If T is A -relatively bounded with relative bound < 1 then $A + T$ defined on the domain $\mathcal{D}(A)$ is closed.*

PROOF. As in the self-adjoint case. □

LECTURE 28

Perturbation of eigenvalues

Throughout this lecture, let H be a Hilbert space and let $A \in \overline{\mathcal{L}}(H)$ be a closed densely defined operator.

PROPOSITION 28.1. *If $\zeta \in \sigma(A)$ is an isolated point, so $\{0 < |z - \zeta| < r\} \subset \rho(A)$ for some $r > 0$, then*

$$Q = \frac{1}{2\pi i} \oint_{|z-\zeta|=\frac{r}{2}} (zI - A)^{-1} dz,$$

is a bounded projection ($Q^2 = Q$). Furthermore

- (1) $\text{ran } Q \subset \mathcal{D}(A)$,
- (2) $AQx = QAx$, for $x \in \mathcal{D}(A)$
- (3) AQ is bounded
- (4) $\sigma(A|_{\text{ran } Q}) = \{\zeta\}$.

PROOF. The proofs that Q is a bounded projection and of 1, 2, and 3 are left as exercises. That $\sigma(A|_{\text{ran } Q}) = \{\zeta\}$ follows since for $w \neq \zeta$ we have

$$(wI - A) \frac{1}{2\pi i} \oint_{|z-\zeta|=\epsilon} \frac{1}{w-z} (zI - A)^{-1} dz = Q$$

with ϵ sufficiently small that w is outside the contour. □

DEFINITION 28.1. Let $\zeta \in \sigma(A)$ be an isolated point. We will say that ζ is a non-degenerate eigenvalue if $\dim \text{ran } Q = 1$.

Note that if $\dim \text{ran } Q = 1$ then $AQ = \zeta Q$ so ζ is indeed an eigenvalue. Likewise if $\dim \text{ran } Q < \infty$ then ζ is an eigenvalue. (If $\dim \text{ran } Q = \infty$ then ζ may not be an eigenvalue — recall the example of the Volterra operator.)

PROPOSITION 28.2. *Let ζ be an isolated non-degenerate eigenvalue of A then ζ^* is an isolated non-degenerate eigenvalue of $A^\dagger$ and if ϕ, ψ are corresponding non-zero eigenvectors,*

$$A\phi = \zeta\phi, \quad A^\dagger\psi = \zeta^*\psi,$$

then $\langle \phi, \psi \rangle \neq 0$.

PROOF. We have already seen that $\sigma(A^\dagger) = \sigma(A)^*$. Thus ζ^* is an isolated point of $\sigma(A^\dagger)$. Shifting and scaling we may assume without loss that $\zeta = 0$ and that $\{0 < |z| \leq 1\} \subset \rho(A)$. Since $\text{ran } Q = \text{span } \phi$, we have

$$Qu = \langle Qu, \phi \rangle \phi = \langle u, Q^\dagger \phi \rangle \phi$$

for all $u \in H$. Let $\psi = Q^\dagger \phi$, so $Q = \langle \cdot, \psi \rangle \phi$. Since $Q \neq 0$, $\psi \neq 0$. Furthermore, since

$$Q^\dagger = -\frac{1}{2\pi i} \oint_{|z|=1} (z^* I - A^\dagger)^{-1} dz^* = \frac{1}{2\pi i} \oint_{|z|=1} (zI - A)^{-1} dz$$

is the Riesz projection of $A^\dagger$, it follows that $\psi \in \mathcal{D}(A^\dagger)$ and

$$\langle u, A^\dagger \psi \rangle = \langle u, Q^\dagger A^\dagger \psi \rangle = \langle Qu, A^\dagger \psi \rangle = \langle AQ, \psi \rangle = 0$$

for all $u \in H$. Thus $A^\dagger \psi = 0$, so 0 is an eigenvalue of $A^\dagger$. Since

$$Q^\dagger = \langle \cdot, \phi \rangle \psi,$$

we see that $\dim \operatorname{ran} Q^\dagger = 1$. Finally,

$$\langle \phi, \psi \rangle = \langle \phi, Q^\dagger \phi \rangle = \langle Q\phi, \phi \rangle = \|\phi\|^2 \neq 0. \quad \square$$

THEOREM 28.3. *Let $A \in \overline{\mathcal{L}}(H)$, H a Hilbert space, and let T be densely defined and A -relatively bounded. Suppose A has an isolated non-degenerate eigenvalue ζ_0 . Then for sufficiently small κ the operator $A + \kappa T$ has a unique isolated non-degenerate eigenvalue $\zeta(\kappa)$ in a neighborhood of ζ_0 and the map $\kappa \mapsto \zeta(\kappa)$ is analytic. Furthermore, if ϕ_0 and ψ_0 are eigenvectors of A and $A^\dagger$,*

$$A\phi_0 = \zeta_0\psi_0, \quad A\psi_0 = \zeta_0^*\psi_0,$$

normalized so that $\langle \phi_0, \psi_0 \rangle = 1$, there is an H -valued analytic map $\kappa \mapsto \phi_\kappa$ such that

$$A\phi_\kappa + \kappa T\phi_\kappa = \zeta(\kappa)\phi_\kappa$$

and

- (1) $\phi_\kappa = \phi_0 + \kappa(\zeta_0 I - A)^{-1}(I - Q)T\phi_0 + O(\kappa^2)$, where Q is the Riesz projection $Q = \langle \cdot, \psi_0 \rangle \phi_0$,
- (2) $\zeta'(0) = \langle T\phi_0, \psi_0 \rangle$, and
- (3) $\zeta''(0) = \langle T(\zeta_0 I - A)^{-1}(I - Q)T\phi_0, \psi_0 \rangle$.

REMARK. Note that $I - Q : \mathcal{D}(A) \rightarrow \mathcal{D}(A)$ and that

$$A(I - Q) = (I - Q)A.$$

EXERCISE 59. Show that $\zeta_0 \notin \sigma(A|_{\operatorname{ran} I - Q})$, and that

$$(\zeta_0 I - A|_{\operatorname{ran} I - Q})^{-1}(I - Q) = \frac{1}{2\pi i} \oint_{|z - \zeta_0| = \frac{r}{2}} \frac{1}{z - \zeta_0} (zI - A)^{-1} dz.$$

PROOF. By shifting and scaling, we may assume that $\zeta_0 = 0$ and $\{0 < |z| < 2\} \subset \rho(A)$. For sufficiently small κ , κT has A -relative bound less than 1, so $A + \kappa T$ is closed. Let $|z| = 1$. Then for

$$\|T(zI - A)^{-1}\| \leq a \|(zI - A)^{-1}\| + b \|A(zI - A)^{-1}\| \leq (a + 1) \max_{|z|=1} \|(zI - A)^{-1}\| + b.$$

So for sufficiently small κ , $z \in \rho(A + \kappa T)$ and

$$(zI - A - \kappa T)^{-1} = \sum_{n=0}^{\infty} \kappa^n (zI - A)^{-1} [T(zI - A)^{-1}]^n$$

is analytic in κ . It follows that the family of Riesz projections

$$Q_\kappa = \frac{1}{2\pi i} \oint_{|z|=1} (zI - A - \kappa T)^{-1} dz$$

is an analytic family of projections.

EXERCISE 60. Show that $\dim Q_\kappa$ is independent of κ . (Hint: look at the proof of the constancy of deficiency indices.)

Thus for small enough κ , Q_κ is a rank one projection and $A + \kappa T$ has an isolated non-degenerate eigenvalue in the disk $|z| < 1$. Let $\zeta(\kappa)$ be the eigenvalue and

$$\phi_\kappa = Q_\kappa \phi_0$$

be the eigenvector. (Note that $Q_\kappa \phi_0 \neq 0$ for small κ .) So ϕ_κ is an analytic family of vectors and

$$(A + \kappa T)\phi_\kappa = \zeta(\kappa)\phi_\kappa,$$

which shows that $\zeta(\kappa)$ is analytic.

The Neumann series gives explicitly

$$\phi_\kappa = \phi_0 + \kappa \phi_1 + \kappa^2 \phi_2 + O(\kappa^3),$$

where

$$\phi_1 = \frac{1}{2\pi i} \oint_{|z|=1} (zI - A)^{-1} T (zI - A)^{-1} \phi_0 = \frac{1}{2\pi i} \oint_{|z|=1} \frac{1}{z} (zI - A)^{-1} T \phi_0 = -A^{-1} (I - Q) T \phi_0,$$

and $\phi_2 \in \mathcal{D}(A)$ could be computed from the second order term. Since $(A + \kappa T)\phi_\kappa = \zeta(\kappa)\phi_\kappa$ it follows that

$$\zeta'(0)\phi_0 = A\phi_1 + T\phi_0 = T\phi_0 - (I - Q)T\phi_0 = QT\phi_0,$$

so

$$\zeta'(0) = \langle T\phi_0, \psi_0 \rangle.$$

Likewise

$$\frac{1}{2}\zeta''(0)\phi_0 = T\phi_1 + A\phi_2 - \zeta'(0)\phi_1.$$

Thus

$$\zeta''(0) = 2\langle T\phi_1, \psi_0 \rangle + 2\langle A\phi_2, \psi_0 \rangle - 2\zeta'(0)\langle \phi_1, \psi_0 \rangle = -2\langle TA^{-1}(I - Q)T\phi_0, \psi_0 \rangle. \quad \square$$