LECTURE NOTES ON THE SPECTRAL THEOREM

DANA P. WILLIAMS

Abstract. Sections 1 through 5 of these notes are from a series of lectures I gave in the summer of 1989. The object of these lectures was to give a reasonably self-contained proof of the Spectral Theorem for bounded normal operators on an infinite dimensional complex Hilbert space. They are aimed at second year graduate students who have at least had a bit of functional analysis. Since the talks, and in particular these notes, were meant to be informal, I did not take the time to carefully footnote the sources for the arguments which I stole from the standard texts. On can safely assume that the clever bits may be found—word for word—in either [8, Chapters 10–13] and/or [1, §1.1].

I added Section 6 since I was curious about non normal operators and what can be said about them. Section 7 was added in 2006 based on a series of talks in our functional analysis seminar. Section 6 comes from [8, §§10.21–33]. Section 7 is based on some old lecture notes from a course I took from my advisor, Marc Rieffel, in 1978 and [8, Chap. 13].

Contents

1. Notation, Assumptions and General Introduction

First let's establish some notation and establish some ground rules.

- Life takes place in complex Hilbert space. (Usually called $\mathcal{H}$.) Consequently, the only scalar field I'll be using will be the complex numbers, $\mathbf{C}$.
- A linear operator $T : \mathcal{H} \to \mathcal{H}$ is said to be bounded if
$$\|T\| = \sup_{|\xi|\leq 1} \|T\xi\| < \infty.$$
- If $\mathcal{H}$ is finite dimensional, then all linear operators are bounded. To see this, note that every operator is given by matrix multiplication; therefore, T is continuous from $\mathcal{H}$ to $\mathcal{H}$. Now you can either prove that (in general)

Date: August 10, 2018.

a linear operator $T : \mathcal{H} \to \mathcal{H}$ is continuous if and only if it is bounded, or you can note that $\{ \xi \in \mathbf{C}^n : |\xi| \le 1 \}$ is compact.

- Every bounded operator T has an adjoint T^* such that $(T\xi, \eta) = (\xi, T^*\eta)$ for all $\xi, \eta \in \mathcal{H}$. (For fixed η, $\xi \mapsto (T\xi, \eta)$ is a bounded linear functional; hence, there exists $T^*\eta \in \mathcal{H}$ so that $(T\xi, \eta) = (\xi, T^*\eta)$.) One can show that T^* is a bounded linear operator with $\|T^*\| = \|T\|$. If $\mathcal{H}$ is finite dimensional, then the matrix of T^* is the conjugate transpose of that of T.
- An operator is called normal if $T^*T = TT^*$. Of course, an operator is called self-adjoint if $T = T^*$. Finally, T is called positive if $(T\xi, \xi) \ge 0$ for all $\xi \in \mathcal{H}$. A positive operator is self-adjoint, and a self-adjoint operator is normal.
- As operators on $\mathbf{C}^2$, $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ is not normal, $\left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$ is normal but not self-adjoint, $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 2 \end{smallmatrix}\right)$ is self-adjoint but not positive. The operator $\left(\begin{smallmatrix} 3 & -2 \\ -2 & 2 \end{smallmatrix}\right)$ is positive.
- If $T \in B(\mathcal{H})$, then the spectrum of T is
$$\sigma(T) = \{ \lambda \in \mathbf{C} : \lambda I - T \text{ is not invertible} \}.$$

Example 1.1. Let X be a non-empty compact subset of $\mathbf{C}$, and μ a finite measure on X with full support (i.e., $\mu(V) > 0$ for every nonempty open set $V \subseteq X$). Let $\mathcal{H} = L^2(X, \mu)$. For each $f \in C(X)$ define M_f by
$$M_f\xi(x) = f(x)\xi(x).$$

(A) Then $M_f \in B(\mathcal{H})$, $\|M_f\| = \|f\|_\infty$, $M_f^* = M_{\bar{f}}$, and $\sigma(M_f) = f(X)$.
(B) Note that $M_gM_f = M_{gf}$. Consequently, M_f is normal. Of course, M_f will be self-adjoint if $f(X) \subseteq \mathbf{R}$, and positive if $f(X) \subseteq [0, \infty)$.
(C) Note that if μ is continuous (i.e., $\mu(\{ x \}) = 0$ for all $x \in X$) and if $f(x) = x$ for all $x \in X$, then M_f *has no eigenvalues*!

I've included the next example more for reference than anything else. It shouldn't be taken too seriously on a first read through and I certainly won't use it elsewhere.

Example 1.2 (Ultimate Example). Now let X be a compact subset of $\mathbf{C}$, and suppose that $X = \bigcup_{n=1}^{n=\infty} X_n$ is a Borel partition of X. Let
$$\mathcal{H}_n = L^2(X_n, \mu_n) \otimes \mathbf{C}^n \cong L^2(X_n, \mu_n, \mathbf{C}^n) \quad (n = 1, 2, \dots, \infty).$$
(Here $\mathbf{C}^\infty$ denotes ℓ^2 or your favorite infinite dimensional separable Hilbert space). Then $\widetilde{\mathcal{H}} = \bigoplus_{n=1}^{n=\infty} \mathcal{H}_n$ is a Hilbert space and M_f is defined in the obvious way for $f \in C(X)$. Assertions (A), (B), and (C) are still valid.

The punch line is that every normal operator on a separable Hilbert space is unitarily equivalent to such a M_f. (In fact, one can take $X = \sigma(T)$ and $f(\lambda) = \lambda$). That is, there is a Hilbert space isomorphism $U : \mathcal{H} \to \widetilde{\mathcal{H}}$ so that
$$T = U^*M_fU$$
This is a pretty, but not particularly useful, abstract version of the spectral theorem.

To motivate what follows, let's review the spectral theorem for $\mathcal{H}$ finite dimensional.

Lemma 1.3. *Suppose that $\mathcal{H}$ is finite dimensional. If $T \in B(\mathcal{H})$, then $\sigma(T) \neq \emptyset$. Furthermore, $\lambda \in \sigma(T)$ if and only if λ is an eigenvalue for T.*

Lemma 1.4. *If $T \in B(\mathcal{H})$ is normal, and if v is a eigenvector for T with eigenvalue λ, then v is a eigenvector for T^* with eigenvalue $\bar{\lambda}$.*

Proof. Let $\mathcal{E}_\lambda = \{ v \in \mathcal{H} : Tv = \lambda v \}$. Since $TT^*v = T^*Tv = \lambda T^*v$, we have $T^*v \in \mathcal{E}_\lambda$. If $w \in \mathcal{E}_\lambda$, then

$$\begin{aligned}(T^*v - \bar{\lambda}v, w) &= (T^*v, w) - \bar{\lambda}(v, w) \\ &= (v, Tw) - \bar{\lambda}(v, w) \\ &= 0.\end{aligned}$$

Since $T^*v - \bar{\lambda}v \in \mathcal{E}_\lambda$, we have $T^*v = \bar{\lambda}v$. □

Proposition 1.5. *Suppose that $\mathcal{H}$ is finite dimensional. If $T \in B(\mathcal{H})$ is normal, then $\mathcal{H}$ has a basis of eigenvectors for T.*

Proof. By Lemmas 1.3 and 1.4, there is a $v \in \mathcal{H}$, which is an eigenvector for both T and T^*. Let $W = \{v\}^\perp$. Then W is invariant under T^* and T. It follows that $T|_W = S$ is a normal operator on $B(W)$. The result follows by induction. □

In most undergraduate texts, Proposition 1.5 is called the *Principal Axes Theorem.* It needs to be dressed up a little before it can enjoy being called the *Spectral Theorem.* Now, still in the finite dimensional case, let $\sigma(T) = \{\lambda_1, \dots, \lambda_n\}$. Let $\mathcal{E}_n = \mathcal{E}_{\lambda_n}$ and P_n the orthogonal projection onto $\mathcal{E}_n$. Standard nonsense implies that $P_nP_m = 0$ if $n \neq m$. Moreover, by Proposition 1.5

$$T = \lambda_1 P_1 + \cdots + \lambda_n P_n. \tag{1.1}$$

Define $\Psi : C\big(\sigma(T)\big) \to B(\mathcal{H})$ by

$$\Psi(f) = f(\lambda_1)P_1 + \cdots + f(\lambda_n)P_n. \tag{1.2}$$

Note that if p is a polynomial, then $\Psi(p) = p(T)$ where the latter has the obvious meaning. Thus, one often writes $f(T)$ for general f.

It is not hard to see that Ψ is a isometric $*$-isomorphism into its range: i.e.,

(a) $\|\Psi(f)\| = \|f\|_\infty$
(b) $\Psi(f + g) = \Psi(f) + \Psi(g)$
(c) $\Psi(fg) = \Psi(f)\Psi(g)$
(d) $\Psi(\bar{f}) = \Psi(f)^*$

(Note that (c) is obvious for polynomials, and every $f \in C\big(\sigma(T)\big)$ is the restriction of a polynomial to $\sigma(T)$). The existence of this isomorphism can be very useful; the process of passing from $f \in C(\sigma(T))$ to $f(T)$ is called the functional calculus. It allows us to construct, for example, square roots, logs, and exponentials of operators. In fact, the functional calculus can be used to construct a great many interesting operators. For this reason, the decomposition of (1.1) is usually called the Spectral Theorem in finite dimensions.

Although, one can generalize (1.2) directly (and we will), the notation used is, unfortunately, often a different one. We introduce it here, so that it won't seem so strange latter.

1.1. **Alternate Notation.** For each $A \subseteq \sigma(T)$, let

$$P(A) = \sum_{\lambda_i \in A} P_i.$$

The map $P : \mathcal{P}\bigl(\sigma(T)\bigr) \to \mathcal{B}(\mathcal{H})$ is a special case of what we will call a "$\mathcal{H}$-projection valued measure" on $\sigma(T)$. (Notice that if $v, w \in \mathcal{H}$, then $\mu_{v,w}(A) = \bigl(P(A)v, w\bigr)$ is a bonafide measure on $\sigma(T)\bigr)$. When using this approach, one writes

$$\int_{\sigma(T)} f(\lambda)\, dP(\lambda) \quad \text{in place of } \Psi(f)$$

This notation is (marginally) justified by the fact that

$$\bigl(\Psi(f)v, w\bigr) = \Bigl(\int_{\sigma(T)} f\, dPv, w\Bigr) = \int_{\sigma(T)} f\, d\mu_{v,w}.$$

In general, the classical Spectral Theorem says that each normal $T \in \mathcal{B}(\mathcal{H})$ is associated to a (unique) $\mathcal{H}$-projection valued measure on $\sigma(T)$, which we will define later, so that

$$T = \int_{\sigma(T)} \lambda\, dP(\lambda).$$

Furthermore, $\lambda \in \sigma(T)$ is an eigenvalue for T if and only if $P(\{\lambda\}) \neq 0$. In the Example 1.2 on page 2, one defines

$$P(E) = M_{\chi_E}$$

for each measurable set $E \subseteq X$. Again, one can check that if $\xi, \eta \in \widetilde{\mathcal{H}}$, then $\mu_{\xi,\eta}(E) = \bigl(P(E)\xi, \eta\bigr) = \int_E \xi(\lambda)\overline{\eta(\lambda)}\, d\mu(\lambda)$ is an honest measure and

$$(M_f \xi, \eta) = \int f(\lambda)\, d\mu_{\xi,\eta}(\lambda).$$

2. The spectrum

Definition 2.1. A Banach algebra is a complex Banach space which is an algebra in such a way that $\|xy\| \leq \|x\|\|y\|$. *Our algebras will always be assumed to have an identity e.*

Example 2.2. Let $A = B(\mathcal{H})$. Then, if $T, S \in A$, it is easy to see that $\|ST\| \leq \|S\|\|T\|$, and that A is a Banach algebra.

Definition 2.3. If A is a Banach algebra, then let $G(A)$ denote the group of invertible elements.

It is a standard exercise to show that if A is a Banach algebra and $x \in A$ with $\|x\| < 1$, then $e - x \in G(A)$. In fact, $(e - x)^{-1} = e + x + x^2 + \cdots$.

Lemma 2.4. *$G(A)$ is open.*

Proof. Suppose $x \in G(A)$. Then $x + h = x(e + x^{-1}h)$, and $(e + x^{-1}h) \in G(A)$ if $\|h\| < \|x^{-1}\|^{-1}$. $\square$

Definition 2.5. Let A be a Banach algebra with unit e. If $x \in A$, then the *spectrum* of x is
$$\sigma(x) = \{\lambda \in \mathbf{C} : \lambda e - x \notin G(A)\}$$
and the *spectral radius* of x is
$$\rho(x) = \sup\{\, |\lambda| \in \mathbf{C} : \lambda \in \sigma(x)\,\}.$$

Theorem 2.6. *If A is a Banach algebra with unit e, then for each $x \in A$,*

(a) (Gelfand) $\sigma(x)$ *is compact and nonempty, and*
(b) (Beurling) $\rho(x) = \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = \inf_{n\geq 1} \|x^n\|^{\frac{1}{n}}$.

Remark 2.7. The existence of the limit is part of the conclusion as is the fact that $\rho(x) \leq \|x\|$.

Proof. If $|\lambda| > \|x\|$, then $\lambda e - x = \lambda(e - \lambda^{-1}x) \in G(A)$. Therefore, $\rho(x) \leq \|x\|$; in particular $\sigma(x)$ is bounded.

Define $g : \mathbf{C} \to A$ by $g(\lambda) = \lambda e - x$. Then g is continuous. Thus,
$$\Omega = \{\, \lambda \in \mathbf{C} : \lambda \notin \sigma(x)\,\} = g^{-1}(G(A))$$
is open. Therefore, $\sigma(x)$ is closed. (Hence, compact!). Note that if $x = 0$, then $\sigma(x) = \{0\}$ and we're done. So we can assume $x \neq 0$ in the sequel.

Now define $f : \Omega \to G(A)$ by $f(\lambda) = (\lambda e - x)^{-1}$. (One often writes $f(\lambda) = R(x, \lambda)$ and refers to $f(\lambda)$ as the resolvent of x at λ).

Now consider
$$\begin{aligned}\lim_{h\to 0} \frac{1}{h}(f(\lambda+h) - f(\lambda)) &= \lim_{h\to 0} \frac{1}{h}((\lambda e - x + he)^{-1} - (\lambda e - x)^{-1}) \\ &= \lim_{h\to 0} \frac{1}{h}\Big((e + f(\lambda)h)^{-1} - e\Big) f(\lambda),\end{aligned}$$
Now if h is so small that $\|f(\lambda)h\| < 1$, then
$$= \lim_{h\to 0} \frac{1}{h}\Big(\sum_{n=1}^{\infty} \big(-f(\lambda)h\big)^n\Big) f(\lambda) = -f(\lambda)^2.$$
In plain terms, $\lambda \mapsto f(\lambda)$ is a strongly holomorphic A-valued function on Ω.[1]

Now if $\lambda > \|x\|$, then
$$\begin{aligned} f(\lambda) = (\lambda e - x)^{-1} &= \lambda^{-1}\Big(e - \frac{x}{\lambda}\Big)^{-1} \\ &= \frac{1}{\lambda}\sum_{n=0}^{\infty}\Big(\frac{x}{\lambda}\Big)^n \\ &= \frac{1}{\lambda}e + \frac{1}{\lambda^2}x + \cdots .\end{aligned}$$

And the convergence is uniform on circles Γ_r centered at 0 provided $r > \|x\|$. Since $\int_{\Gamma_r} \lambda^n \, d\lambda$ is $2\pi i$ if n is -1 and 0 otherwise, we get:[2]

---

[1]Alternatively, $\lambda \mapsto \Lambda(f(\lambda))$ is (honestly) holomorphic on Ω for every $\Lambda \in A^*$.

[2]Again, one may apply a $\Lambda \in A^*$ to both sides so that
$$\Lambda(x^n) = \frac{1}{2\pi i}\int_{\Gamma_r} \lambda^n \Lambda(f(\lambda))\, d\lambda$$
for all $\Lambda \in A^*$.

$$x^n = \frac{1}{2\pi i} \int_{\Gamma_r} \lambda^n f(\lambda)\, d\lambda \qquad n = 0, 1, 2, \dots \tag{2.1}$$

Since $|\lambda| > \rho(x)$ implies that $\lambda \in \Omega$, it follows that if $r, r' > \rho(x)$, then Γ_r and $\Gamma_{r'}$ are *homotopic* in Ω! Thus, Equation 2.1 holds for all $r > \rho(x)$.

Now if $\Omega = \mathbf{C}$, then f is entire and bounded (since $|\lambda| > 2\|x\|$ implies that $\|f(\lambda)\| \leq \frac{1}{|\lambda|} \frac{1}{1 - \frac{\|x\|}{\lambda}} \leq \frac{1}{2\|x\|} \frac{1}{1-\frac{1}{2}} = \frac{1}{\|x\|}$). Therefore, f would be constant. It would then follow from Equation 2.1, with $n = 0$, that $e = 0$; this is a contradiction, so $\sigma(x) \neq \emptyset$. We've proved (a).

Now let $M(r) = \max_{\lambda \in \Gamma_r} \|f(\lambda)\|$. Using Equation 2.1,

$$\|x^n\| \leq r^{n+1} M(r).$$

Thus,

$$\limsup_n \|x^n\|^{\frac{1}{n}} \leq r.$$

It follows that

$$\limsup_n \|x^n\|^{\frac{1}{n}} \leq \rho(x).$$

Now, if $\lambda \in \sigma(x)$, then

$$\begin{aligned}(\lambda^n e - x^n) &= (\lambda e - x)(\lambda^{n-1} e + \lambda^{n-2} x + \cdots + x^{n-1}) \\ &= (\lambda^{n-1} e + \lambda^{n-2} x + \cdots + x^{n-1})(\lambda e - x)\end{aligned}$$

implies that $\lambda^n \in \sigma(x^n)$—otherwise, $(\lambda^{n-1}e + \cdots + x^{n-1})(\lambda^n e - x^n)^{-1}$ is an inverse for $(\lambda e - x)$. As a consequence, if $\lambda \in \sigma(x)$, then $|\lambda^n| \leq \rho(x^n) \leq \|x^n\|$. In particular, $|\lambda| \leq \|x^n\|^{\frac{1}{n}}$. Therefore, $\rho(x) \leq \inf_{n\geq 1} \|x^n\|^{\frac{1}{n}}$. Combining with the previous paragraph,

$$\limsup_n \|x^n\|^{\frac{1}{n}} \leq \rho(x) \leq \inf_{n \geq 1} \|x^n\|^{\frac{1}{n}} \leq \liminf_n \|x^n\|^{\frac{1}{n}}.$$

This completes the proof. □

Corollary 2.8 (Gelfand-Mazur)**.** *A unital Banach algebra in which every non-zero element is invertible is isometrically isomorphic to* $\mathbf{C}$*.*

Proof. If $\lambda \neq \mu$, then at most one of $\lambda e - x$ and $\mu e - x$ can be zero for any $x \in A$. Thus, $\sigma(x)$ consists of a single number—say $\sigma(x) = \{\, \lambda(x) \,\}$. By assumption $\lambda(x)e - x = 0$. One can prove that $\lambda : A \to \mathbf{C}$ is the required map[3]. □

---

[3]For example,

$$\begin{aligned}\lambda(x)\lambda(y)e - xy &= \lambda(x)\lambda(y)e - \lambda(x)y + \lambda(x)y - xy \\ &= \lambda(x)(\lambda(y)e - y) + (\lambda(x) - x)y = 0.\end{aligned}$$

Thus, $\lambda(xy) = \lambda(x)\lambda(y)$.

3. The Gelfand transform

Recall that $J \subsetneqq A$ is called a maximal ideal if J is a proper ideal which is contained in no larger *proper* ideal. Since $G(A)$ is open and any proper ideal satisfies $J \cap G(A) = \emptyset$, it follows that maximal ideals are closed – given any proper ideal, its closure is an ideal disjoint from $G(A)$ and hence proper. A linear functional $h : A \to \mathbf{C}$ which is also multiplicative is called a *complex homomorphism.*

Theorem 3.1. *Let A be a unital commutative Banach algebra. Let Δ denote the collection of nonzero complex homomorphisms.*

(1) J is a maximal ideal in A if and only if J is the kernel of some $h \in \Delta$.
(2) $\|h\| = 1$ for all $h \in \Delta$.
(3) $\lambda \in \sigma(x)$ if and only if $h(x) = \lambda$ for some $h \in \Delta$.

Proof. Let J be a maximal ideal in A. Since J is closed, the natural map $\pi : A \to A/J$ has norm 1.[4] If $x \in A$ is such that $\pi(x) \neq 0$ — so that $x \notin J$ — then let

$$M = \{\, ax + y : a \in A \text{ and } y \in J \,\}.$$

Then M is an ideal in A such that $J \subsetneqq M$. Therefore, $M = A$; in particular, for some $a \in A$ and $y \in J$,

$$ax + y = e.$$

Hence, $\pi(a)\pi(x) = \pi(e)$. Thus, $\pi(x)$ is invertible. It follows from Corollary 2.8 that A/J is isomorphic to $\mathbf{C}$, and that π defines a complex homomorphism with kernel J. This establishes the first part of part (1).

On the other hand, if $h \in \Delta$, then we must show that $h^{-1}(0)$ is a maximal ideal. However, if $J \supseteq h^{-1}(0)$, and if there is a $x \in J$ with $h(x) \neq 0$, then, given $y \in A$, we have $y - h(y)h(x)^{-1}x \in h^{-1}(0) \subseteq J$. It follows that $y \in J$, and that $J = A$. This proves (1).

Part (2) follows from (1) and the fact that quotient maps have norm 1.

To prove (3), first notice that $x \in G(A)$ implies $h(x)h(x^{-1}) = h(e) = 1$ for all $h \in \Delta$. Thus, $x \in G(A)$ implies that $h(x) \neq 0$ for all $h \in \Delta$. On the other hand, if $x \notin G(A)$, then $J = \{\, ax : a \in A \,\}$ is a *proper* ideal. Since J is contained in a maximal ideal, $J \subseteq \ker h$ for some $h \in \Delta$ by part (1). Thus, $x \in G(A)$ if and only if $h(x) \neq 0$ for all $h \in \Delta$. The result follows by replacing x by $\lambda e - x$ above. ☐

Definition 3.2. Let A be a unital commutative Banach algebra and Δ the collection of nonzero complex homomorphisms of A. For each $x \in A$, the *Gelfand transform* of x is the function $\hat{x} : \Delta \to \mathbf{C}$ defined by $\hat{x}(h) = h(x)$. The Gelfand topology on Δ is the smallest topology for which each $\hat{x}$ is continuous. The set $\Delta = \Delta(A)$ equipped with the Gelfand topology is called the *maximal ideal space* of A, or the *spectrum* of A.

We need a brief digression into “point-set topology”.

Definition 3.3. Let X be a set, Y be a topological space and $\mathcal{F}$ a collection of functions $f : X \to Y$. Then the *initial topology* on X is the smallest topology such that each $f \in \mathcal{F}$ is continuous.

A subbasis for the initial topology on X is given by

$$\rho = \{\, f^{-1}(U) : U \subset Y \text{ is open and } f \in \mathcal{F} \,\}.$$

[4]We need J closed so that A/J is a normed space—a Banach algebra in fact—$\|\pi(x)\| = \inf_{y \in J} \|x + y\|$.

Example 3.4. The weak-$*$ topology on $X = A^*$ is the initial topology for $\mathcal{F} := \{ a \in A \}$ where each $a \in A$ is viewed a complex-valued function $(\varphi \mapsto \varphi(a))$ on A^*.

Lemma 3.5. *Suppose that X has the initial topology corresponding to $\mathcal{F} \subset X^Y$ for some space Y. If $S \subset X$, then the relative topology on S is the initial topology for $\mathcal{F}|_S := \{ f|_S : f \in \mathcal{F} \}$.*

Proof. ${f|_S}^{-1}(U) = f^{-1}(U) \cap S$. □

Theorem 3.6. *Let A be a unital commutative Banach algebra with maximal ideal space Δ. Then Δ is a compact Hausdorff space (with the Gelfand topology), and the Gelfand transform is a homomorphism of A into $C(\Delta)$ with kernel*

$$\operatorname{rad}(A) = \bigcap \{ J : J \text{ is a maximal ideal in } A \}.$$

Moreover, $\|\hat{x}\|_\infty = \rho(x)$.

Remark 3.7. If $\operatorname{rad}(A) = \{ 0 \}$, then A is called semisimple. Consequently, the Gelfand transform is injective if and only if A is semisimple.

Proof. We have $\hat{x} \in C(\Delta)$ by definition, and it is straightforward to check that $x \mapsto \hat{x}$ is an algebraic homomorphism. Since $\hat{x} = 0$ if and only if $h(x) = 0$ for all $h \in \Delta$, the formula for $\operatorname{rad}(A)$ follows from part (1) of Theorem 3.1 and, the formula for $\|\hat{x}\|_\infty$ follows from part (3) of the same theorem. Therefore we only have to prove that Δ is compact and Hausdorff.

However, $K = \{ \varphi \in A^* : \|\varphi\| \leq 1 \}$ is weak-$*$ compact by the Banach-Alaoglu theorem,[5] and $\Delta \subseteq K$. It follows from Example 3.4 and Lemma 3.5 that the Gelfand topology is the restriction of the weak-$*$ topology to Δ. We'll be done once we see that Δ is closed in K (since the weak-$*$ topology is Hausdorff).

Now let h_α be a net converging to φ in K. Then $h_\alpha(x) \to \varphi(x)$ for all $x \in A$. Thus, $\varphi(e) = 1$ and $\varphi(xy) = \lim_\alpha h_\alpha(xy) = \varphi(x)\varphi(y)$. In short, $\varphi \in \Delta$. □

Since we are eventually going to want to concentrate on subalgebras of $B(\mathcal{H})$, we want to start to pay attention to the fact that operators on Hilbert space have adjoints. (The adjoint *has* to play a significant rôle since, even in finite dimensions, only normal operators have a decent spectral decomposition!) Therefore we make the following definition.

Definition 3.8. A Banach $*$-algebra is a Banach algebra A together with a map $x \mapsto x^*$ of A onto itself which satisfies

$$\begin{aligned} &\text{(a)} \quad x^{**} = x \\ &\text{(b)} \quad (xy)^* = y^* x^* \\ &\text{(c)} \quad (x + \lambda y)^* = x^* + \bar{\lambda} y^* \\ &\text{(d)} \quad \|x^*\| = \|x\| \end{aligned}$$

for all $x, y \in A$ and $\lambda \in \mathbf{C}$.

Example 3.9. Of course, the motivating example is $B(\mathcal{H})$, or any self-adjoint subalgebra of $B(\mathcal{H})$.

[5]Let $D_r = \{ \zeta \in \mathbf{C} : |\zeta| \leq r \}$. Then $C = \prod_{a \in A} D_{\|a\|}$ is compact by Tychonoff's Theorem. Then you check that the map $\kappa : K \to C$ defined by $\kappa(\varphi)(a) = \varphi(a)$ is a homeomorphism onto a closed subset of C.

Example 3.10. Let G be a locally compact abelian group with Haar measure λ. Then by applying Fubini's theorem and some standard measure theory we get that $A = L^1(G, \lambda)$ is a commutative Banach $*$-algebra with multiplication

$$f * g(s) = \int_G f(t)g(s-t)\,d\lambda(t),$$

and involution

$$f^*(s) = \overline{f(-s)}.$$

Unfortunately, A is unital if and *only if* G is discrete. However, it is not hard to see that if G is not discrete, then

$$B = \mathbf{C} \oplus A = \{\, (\lambda, f) \in \mathbf{C} \times A \,\}$$

is a unital commutative Banach $*$-algebra when one defines

$$\begin{aligned} &(1) \quad \|(\lambda, f)\| = |\lambda| + \|f\|_1 \\ &(2) \quad (\lambda, f)(\mu, g) = (\lambda\mu, \lambda g + \mu f + f * g) \\ &(3) \quad (\lambda, f)^* = (\bar{\lambda}, f^*) \end{aligned}$$

Of course, in order to check the norm inequality in Definition 2.1 you will want to recall the fact that $\|f * g\|_1 \le \|f\|_1 \|g\|_1$.

Furthermore, if $w \in \widehat{G}$, then

$$\begin{aligned} h_w((\lambda, f)) &= \lambda + \int_G f(t)\overline{w(t)}\,dt \\ &= \lambda + \hat{f}(w) \end{aligned}$$

can be shown to be a complex homomorphism (here $\widehat{G}$ is the group of unitary characters on G, and $\hat{f}$ is the Fourier transform). In fact, although this requires some work, every $h \in \Delta(B)$ is either of the form h_w for some $w \in \widehat{G}$, or $h((\lambda, f)) = \lambda$. Remarkably, $\Delta(B)$ is the one point compactification of $\widehat{G}$ (with the dual topology from abelian harmonic analysis—the topology of uniform convergence on compacta). The point is that the Fourier transform is a special case of the Gelfand transform.

4. The Abstract Spectral Theorem

Bounded operators on Hilbert space have one more special property we require:

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

This is the (or at least a) reason for the next definition.

Definition 4.1. A Banach $*$-algebra A is called a C^*-algebra if $\|x^*x\| = \|x\|^2$ for all $x \in A$.

Example 4.2. Let A be a self-adjoint, norm closed subalgebra of $B(\mathcal{H})$. Then A is a C^*-algebra. In particular, if $T \in B(\mathcal{H})$ is normal—that is, $TT^* = T^*T$—then the closure in the norm topology of the algebra generated by I, T, and T^* is a unital commutative C^*-algebra.

It is worth noting that the algebras in Example 3.10 are (almost) never C^*-algebras.

Example 4.3. Suppose that X is a compact Hausdorff space. Then $A = C(X)$ is a unital commutative C^*-algebra with maximal ideal space (homeomorphic to) X. In particular, every complex homomorphism is of the form $h_x : A \to \mathbf{C}$, where $h_x(f) = f(x)$.

Proof. A is obviously a Banach $*$-algebra with multiplication defined by $fg(x) = f(x)g(x)$ and with $f^*(x) = \bar{f}(x) = \overline{f(x)}$. Moreover, A is a C^*-algebra as

$$\|f^*f\|_\infty = \||f|^2\|_\infty = \|f\|_\infty^2.$$

Now suppose that J is a closed ideal in A with the property that given $x \in X$, there is a $f \in J$ such that $f(x) \neq 0$. Then using the compactness of X, there are $f_1, f_2, \dots, f_n \in J$ so that $\sum_i |f_i|^2$ does not vanish on X. Then $1 = (\sum_i |f_i|^2)^{-1} \sum_i |f_i|^2$ must belong to J. Of course that means $J = A$. It follows that every maximal ideal is of the form $J = \{ f \in A : f(x) = 0 \}$ for some $x \in X$. In particular, every $h \in \Delta$ is a point evaluation, since $h(g) = h(g - g(x) \cdot 1 + g(x) \cdot 1) = h(g - g(x)1) + g(x) \cdot h(1) = g(x)$ since $g - g(x)1 \in J$. Finally, it is not hard to see that the map $\rho : x \mapsto h_x$ is continuous from X to Δ (Δ has initial topology induced by $\hat{x}$, and for all $\theta \in A, \hat{\theta} \circ \rho = \theta$ is continuous) as well as one-to-one and onto. Since X is compact and Δ is Hausdorff, they must be homeomorphic. □

Definition 4.4. If A and B are Banach $*$-algebras then a homomorphism $\Phi : A \to B$ is called a $*$-homomorphism if $\Phi(a^*) = \Phi(a)^*$ for all $a \in A$.

Remarkably, Example 4.2 "includes" all C^*-algebras in the sense that every C^*-algebra is isometrically $*$-isomorphic to such an algebra. Unfortunately, we don't need this fact, so I can't justify proving it.[6] The Abstract Spectral Theorem is, however, merely a characterization of (unital) commutative C^*-algebras which says that Example 4.3 is the "only" example.

Theorem 4.5 (Abstract Spectral Theorem)**.** *Suppose that A is a unital commutative C^*-algebra. Then the Gelfand map is an isometric $*$-isomorphism of A onto $C(\Delta)$.*

Remark 4.6. The word "onto" in the statement is critical here. Notice also that the "$*$"-part simply means that $\widehat{x^*}(h) = \overline{\hat{x}(h)}$. This result should be compared with Equation 1.2.

Proof. Let $h \in \Delta$. We first want to show that $h(x^*) = \overline{h(x)}$ (thus, $\widehat{x^*} = \bar{\hat{x}}$). Since each $x \in A$ equals $\frac{x+x^*}{2} - i\frac{ix+(ix)^*}{2}$, each $x \in A$ can be written as $x_1 + ix_2$ with $x_i^* = x_i$. Thus, it suffices to show that $h(x) \in \mathbf{R}$ if $x = x^*$.

Towards this end, let $t \in \mathbf{R}$. Define

$$u_t = \exp(itx) = \sum_{n=0}^{\infty} \frac{(itx)^n}{n!}.$$

(Notice that $\exp(x + y) = \exp(x)\exp(y)$ if $x, y \in A$, since x and y commute.) Observe that $u_t^* = \exp(-itx)$. Thus,

$$\|u_t\|^2 = \|u_t^* u_t\| = \|\exp(-itx + itx)\| = \|\exp(0)\| = 1.$$

[6] For a proof, see [5, Theorem A.11].

Since $\|h\| = 1$,

$$\begin{aligned}\exp(t\,\mathrm{Re}(ih(x))) &= \exp(\mathrm{Re}(ith(x))) \\ &= \big|\exp(ith(x))\big| \\ &= |h(u_t)| \le 1\end{aligned}$$

for *all* $t \in \mathbf{R}$. Therefore, $h(x) \in \mathbf{R}$.

Next we show that $\|\hat{x}\|_\infty = \|x\|$. But $\|\hat{x}\|_\infty = \rho(x) = \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}$ by Theorem 2.6 and Theorem 3.6. But if $x = x^*$, then

$$\begin{aligned}\|x\|^2 &= \|x^*x\| = \|x^2\|, \text{ and} \\ \|x\|^{2^n} &= \|x^{2^n}\|\end{aligned}$$

by induction. Thus, $\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = \|x\|$ (the limit exists by Theorem 2.6, so checking a subsequence is enough). Thus, $\|\hat{x}\|_\infty = \|x\|$ whenever $x = x^*$. In general,

$$\|\hat{x}\|_\infty^2 = \|\bar{\hat{x}}\hat{x}\|_\infty = \|x^*x\| = \|x\|^2.$$

We have shown that $x \mapsto \hat{x}$ is an isometric $*$-isomorphism of A onto a necessarily *closed*[7] subalgebra of $C(\Delta)$. But the functions in the image clearly separate points of Δ. Furthermore, the image is closed under conjugation, and contains the constant functions (since A is unital). The conclusion follows from the Stone-Weierstrass theorem. □

We have one more technicality to overcome before we can be satisfied with our abstract version of the Spectral Theorem. Namely, if $T \in B(\mathcal{H})$ and A is a C^*-subalgebra of $B(\mathcal{H})$ which contains T, then we have two notions of the spectrum of T: one as an element of $B(\mathcal{H})$ and one as an element of A. In general, if A is an unital Banach algebra and B is a subalgebra with

$$e \in B \subseteq A.$$

Then one clearly has

$$\sigma_A(x) \subseteq \sigma_B(x). \tag{4.1}$$

Unfortunately, as the next example shows, it may happen that the inclusion in (4.1) is proper.

Example 4.7. Let $\mathbf{D} := \{\, z \in \mathbf{C} : |z| < 1 \,\}$ be the open unit disk. Let $A(\mathbf{D})$ be the set of holomorphic functions on $\mathbf{D}$ which have a continuous extension to the boundary $\mathbf{T} = \{\, z \in \mathbf{C} : |z| = 1 \,\}$. By the maximum modulus principle, each $f \in A(\mathbf{D})$ is uniquely determined by its values on $\mathbf{T}$. Therefore, we can view $A(\mathbf{D})$ as a Banach subalgebra of $C(\mathbf{T})$ with respect to the sup norm $\|\cdot\|_\infty$. The reflection principle assures that if $f \in A(\mathbf{D})$, then so is f^* where we define

$$f^*(z) := \overline{f(\bar{z})}. \tag{4.2}$$

Then $A(\mathbf{D})$ is a Banach $*$-subalgebra of $C(\mathbf{T})$ (where both have the involution defined by (4.2)).[8] Now let f be the identity function $z \mapsto z$. Then as an element

[7]Since the map is isometric, the image of A is complete and therefore closed.

[8]By considering $f(z) := i+z$, it is not hard to verify that neither $C(\mathbf{T})$ nor $A(\mathbf{D})$ is a C^*-algebra with this $*$-algebra structure. They are, of course, Banach $*$-algebras.

of $A(\mathbf{D})$, $\lambda - f$ is invertible if and only if $\lambda \notin \overline{\mathbf{D}}$. Thus

$$\sigma_{A(\mathbf{D})}(f) = \mathbf{D} \cup \mathbf{T} \quad \text{while} \quad \sigma_{C(\mathbf{T})}(f) = \mathbf{T}.^{9}$$

On the other hand, if A and B are C^*-algebras, then we can invoke the Abstract Spectral Theorem to show that we always have equality in (4.1).

Theorem 4.8 (Spectral Permanence)**.** *Suppose that B is a unital C^*-subalgebra of a C^*-algebra A (i.e., $e \in B \subseteq A$). Then for all $x \in B$, $\sigma_B(x) = \sigma_A(x)$.*

Proof. Fix $x \in B$. We need only show that $\sigma_B(x) \subseteq \sigma_A(x)$. A moment's reflection shows that it suffices to show that $x \in G(A)$ implies that $x^{-1} \in B$. Furthermore, I claim it suffices to do this for x self-adjoint. To see this notice that $x \in G(A)$ implies that $x^* \in G(A)$, and hence $x^*x \in G(A)$. Thus, $(x^*x)^{-1}x^*$ is a left inverse for x, and, since x^{-1} exists, $x^{-1} = (x^*x)^{-1}x^*$. Thus, it suffices to show that $(x^*x)^{-1} \in B$ if x is; the claim follows.

Let C be the C^*-subalgebra of A generated by x and x^{-1}, and let D be the $*$-subalgebra generated by e and x. Since $(x^{-1})^* = (x^*)^{-1} = x^{-1}$, C is commutative.[10] Thus, $C \cong C(\Delta)$, and $C(\Delta)$ is generated by the functions $\hat{x}$ and $\hat{y} = 1/\hat{x}$ (since $h(x)h(x^{-1}) = 1$ for all $h \in \Delta$). But the image of D in $C(\Delta)$ is generated by 1 and $\hat{x}$. On the other hand, if $\hat{x}(h) = \hat{x}(h')$, then $\hat{y}(h) = \hat{y}(h')$. It follows that $\hat{x}$ must separate points of Δ! Therefore, $D = C$ by the Stone-Weierstrass Theorem. In particular, $x^{-1} \in D \subseteq B$. □

Thus, we see that we may speak unambiguously about the spectrum of an operator $T \in B(\mathcal{H})$ provided we agree to always compute $\sigma(T)$ with respect to some C^*-subalgebra.

Example 4.9 (The functional calculus). Let T be a normal operator in $B(\mathcal{H})$. Let A be the C^*-algebra generated by I and T (and T^*!). By Theorem 4.5, A is isomorphic to $C(\Delta)$ via the Gelfand map. But $\widehat{T} : \Delta \to \mathbf{C}$ is a continuous and one-to-one map of Δ onto $\sigma_A(T) = \sigma(T)$. Since Δ is compact, this map is a homeomorphism. In summary, we have produced a $*$-isomorphism of $C\big(\sigma(T)\big)$ onto A which takes the function $\lambda \mapsto \lambda$ to T. If $f \in C\big(\sigma(T)\big)$, we denote the image of f under this isomorphism by $f(T)$. (Compare this with the discussion following Proposition 1.5 on page 3.)

Remark 4.10 (Composition in the functional calculus). If T is normal and $f \in C\big(\sigma(T)\big)$, then $f(T)$ is normal (its adjoint is just $\bar{f}(T)$), and by spectral permanence (Theorem 4.8), $\sigma\big(f(T)\big) = \sigma_B\big(f(T)\big)$, where $B = C^*(I,T)$. But the functional calculus gives us an isomorphism $\Phi : C\big(\sigma(T)\big) \to B$ such that $\Phi(f) = f(T)$. Thus, $\sigma_B\big(f(T)\big)$ is just the range, $f\big(\sigma(T)\big)$, of f. Therefore spectral permanence gives a baby version of the Spectral Mapping Theorem: $\sigma\big(f(T)\big) = f\big(\sigma(T)\big)$.

Thus if $g \in C\big(f\big(\sigma(T)\big)\big)$, then the functional calculus allows us to define $g\big(f(T)\big)$. Naturally, we expect that $g\big(f(T)\big) = h(T)$, where $h := g \circ f$. Since Φ is an algebra

---

[9] Note that $\sigma_{A(\mathbf{D})}(f)$ is the union of $\sigma_{C(\mathbf{T})}(f)$ and the only bounded connected component of the complement of $\sigma_{C(\mathbf{T})}(f)$. This is a general phenomenon as illustrated by [8, Theorem 10.18].

[10] C is fairly clearly the closure of the commutative algebra of (two variable) polynomials in x and x^{-1}.

homomorphism, this is clear if g is a polynomial in λ and $\bar{\lambda}$.[11] But the Stone-Weierstrass Theorem implies such polynomials are uniformly dense in $C\big(f\big(\sigma(T)\big)\big)$. Thus there are polynomials $g_n \to g$ uniformly on $f\big(\sigma(T)\big)$. Then

$$g_n\big(f(T)\big) \to g\big(f(T)\big).$$

On the other hand, $g_n \circ f \to h := g \circ f$ uniformly on $\sigma(T)$. Therefore,

$$g_n\big(f(T)\big) = (g_n \circ f)(T) \to h(T),$$

and $g\big(f(T)\big) = h(T)$ as we wanted.

Now I can give some examples which hint at the power of Theorem 4.5. The first says that although elements of the spectrum of a *normal* operator T need not be eigenvalues, T does possess "approximate eigenvectors".

Corollary 4.11. *Suppose that T is a normal operator on $\mathcal{H}$, and that $\lambda \in \sigma(T)$. Then there is a sequence $\{\, \xi_n \,\}$ of unit vectors in $\mathcal{H}$ so that $(T - \lambda I)\xi_n$ converges to zero in $\mathcal{H}$.*

Proof. It suffices to consider the case where $\lambda = 0$ (replace T by $T - \lambda I$). By Theorem 4.5 and the functional calculus, there is a $*$-isomorphism $\Psi : C\big(\sigma(T)\big) \to B(\mathcal{H})$ which takes the function defined by $f(\zeta) = \zeta$ to T. For each n, let f_n be a function in $C\big(\sigma(T)\big)$ of norm 1 which equals 1 at 0 (recall we're assuming $0 \in \sigma(T)$), and which vanishes off the disk of radius $1/n$ centered at 0. Notice that $ff_n \to 0$ in norm in $C\big(\sigma(T)\big)$. Since Ψ is isometric, $\Psi(ff_n) = \Psi(f)\Psi(f_n) \to 0$ in $B(\mathcal{H})$! On the other hand, $\|\Psi(f_n)\| = 1$. Thus, we may choose $\eta_n \in \mathcal{H}$ so that $\xi_n = \Psi(f_n)\eta_n$ has norm 1, while $|\eta_n| \leq 2$. Since $T(\xi_n) = \Psi(f)(\xi_n) = \Psi(ff_n)(\eta_n)$, the result follows. $\square$

Corollary 4.12. *Suppose that T is a normal operator in $B(\mathcal{H})$. Then the following statements are equivalent.*

(1) T is a positive operator in $B(\mathcal{H})$.
(2) $\sigma(T) \subseteq [0, \infty)$.
(3) There is an operator $R \in B(\mathcal{H})$ such that $T = R^*R$.

Moreover, T has a unique positive square root $S \in B(\mathcal{H})$ (i.e., $S^2 = T$), and S can be approximated arbitrarily closely in norm by polynomials in T.

Proof. Suppose $\lambda \in \sigma(T)$. Choose $\{\, \xi_n \,\}$ as in Corollary 4.11. Then $((T - \lambda I)\xi_n, \xi_n) \to 0$ implies that $(T\xi_n, \xi_n) \to \lambda$. Since T is positive, it follows that $\sigma(T) \subseteq [0, \infty)$. Therefore, (1) implies (2).

On the other hand, if $\sigma(T) \subseteq [0, \infty)$, then $f(\zeta) = \sqrt{\zeta}$ defines an element of $C(\sigma(T))$. Define S to be $f(T)$ (as in Example 4.9). Now $T = S^2$, and $S^* = \bar{f}(T) = f(T) = S$; we've shown that (2) implies (3).

That (3) implies (1) is immediate.

The existence of a positive square root follows from the above; in fact the S constructed above is positive since $S = R^2$ for $R = \sqrt[4]{T}$. The statement about polynomial approximation follows from the fact that f can be uniformly approximated by polynomials on $\sigma(T)$.

Now let S' be another positive square root. Since S' commutes with T, it commutes with any polynomial in T. Hence, S' must commute with S in view

[11] For example, suppose that $g(\lambda) = \lambda^n \bar{\lambda}^m$. Then $g\big(f(T)\big) = f(T)^n (f(T)^*)^m = f^n(T)\bar{f}^m(G) = g \circ f(T)$.

of the previous paragraph. Thus, $B := C^*(\{ I, S, S' \})$ is commutative (say $\cong C(\Delta)$). Notice that $\widehat{S}$ and $\widehat{S}'$ are nonnegative functions on Δ (for example using Theorem 4.8, $\widehat{S}(\Delta) = \sigma_B(S) = \sigma(S) \subseteq [0, \infty)$). Since S and S' have the same square, $S = S'$. □

Now a meaty example: recall that a (unitary) representation of a locally compact group G is merely a homomorphism π of G into the unitary group of $B(\mathcal{H})$. We insist that the homomorphism be continuous in the sense that $g \mapsto \pi(g)\xi$ should be continuous for all $\xi \in \mathcal{H}$.

Corollary 4.13. *Let G be a locally compact abelian group. Then every irreducible representation of G is one dimensional (i.e., a character).*

Proof. Let $\pi : G \to B(\mathcal{H})$ be an irreducible representation of G. Then

$$\widetilde{\pi}(f) = \int_G f(t)\pi(t)\, d\lambda(t)$$

defines a $*$-homomorphism of $L^1(G, \lambda)$ into $B(\mathcal{H})$ [12]. We may extend $\widetilde{\pi}$ to $B = \mathbf{C} \oplus L1(G, \lambda)$ in the obvious way: $\widetilde{\pi}((\lambda, f)) = \lambda I + \widetilde{\pi}(f)$. The point is that if π has no nontrivial invariant subspaces, then one can show that $\widetilde{\pi}$ has no nontrivial (closed) invariant subspaces [13]. Therefore, $A = \overline{\widetilde{\pi}(B)}$ is a unital commutative C^*-algebra, and A has no nontrivial (closed) invariant subspaces either. If $\Delta = \Delta(A)$ consists of a single point, the result follows easily. If not, then since $A \cong C(\Delta)$, A would contain a closed proper ideal J and a element $x \neq 0$ such that $xJ = \{ 0 \}$. Since J is an ideal, $V = \overline{\operatorname{span}\{ \widetilde{\pi}(y)\xi : y \in J \text{ and } \xi \in \mathcal{H} \}}$ is a nonzero closed invariant subspace of $\mathcal{H}$. However, any vector of the form $\widetilde{\pi}(x)\eta$ for $\eta \in \mathcal{H}$ belongs to $V^\perp$. Since $x \neq 0$, $V^\perp \neq 0$. This is a contradiction. □

Definition 4.14. An operator $K \in B(\mathcal{H})$ is said to be *compact* if

$$K(\mathcal{B}_1) = \{ K\xi \in \mathcal{H} : \xi \in \mathcal{H} \text{ and } |\xi| \leq 1 \} \tag{4.3}$$

has compact closure in $\mathcal{H}$. (Equivalently[14], the image of every bounded sequence has a convergent subsequence.)

Lemma 4.15. *Suppose that K is a bounded normal compact operator in $B(\mathcal{H})$.*

(1) Every nonzero[15] $\lambda \in \sigma(K)$ is an eigenvalue for K.
(2) The eigenspace $\mathcal{E}_\lambda$ is finite dimensional if $\lambda \neq 0$.
(3) $\sigma(K)$ is countable. Furthermore, the only possible accumulation point of $\sigma(K)$ is 0.

---

[12]This integral is Banach space valued. It can be interpreted weakly as saying that

$$(\widetilde{\pi}(f)\xi, \eta) = \int_G f(t)(\pi(t)\xi, \eta)\, d\lambda(t).$$

Note that $\widetilde{\pi}(f)$ is bounded since for every $B : \mathcal{H} \times \mathcal{H} \to C$ bilinear, $\|B(h, k)\| \leq M\|h\|\|k\|$ for some constant M.

[13]If $\{ f_\alpha \}$ is an approximate identity in $L^1(G, \lambda)$,and if we define $s \cdot f_\alpha(t) = f_\alpha(t - s)$, then $\widetilde{\pi}(s \cdot f_\alpha)\xi \to \pi(s)\xi$ for all $\xi \in \mathcal{H}$ and $s \in G$. The assertion follows.

[14]Actually, one can prove that $K(\mathcal{H}_1)$ will be compact whenever it's relatively compact. Furthermore, K will be a compact operator if and only if K is the norm limit of finite rank operators. It follows that the set of compact operators is a norm closed self-adjoint ideal in $B(\mathcal{H})$. These facts aren't particularly difficult to show, but we shall save that path for another day.

[15]In the infinite dimensional case, one always has $0 \in \sigma(K)$ if K is compact. However, it is possible that 0 may not be an eigenvalue—see Example 4.17.

Proof. Let $\lambda \in \sigma(K)$. Choose $\{ \xi_n \}$ as in Corollary 4.11. We may assume that $K\xi_n \to \eta$ in $\mathcal{H}$. Since $(K - \lambda I)\xi_n = K\xi_n - \lambda\xi_n \to 0$, it follows that $\lambda\xi_n \to \eta$ in $\mathcal{H}$ as well. Furthermore, $|\eta| = |\lambda|$ and $K\eta = \lambda\eta$. This proves (1).

Parts (2) and (3) follow from the fact that the image of any orthonormal basis of eigenvectors with eigenvalues bounded away from zero can't have an accumulation point. □

Theorem 4.16 (Spectral Theorem for Compact Operators)**.** *Let K be a normal compact operator in $B(\mathcal{H})$, and let $\{ \lambda_n \}_{n\in I}$ be the nonzero eigenvalues of K. Then, if P_n is the projection*[16] *onto the λ_n-eigenspace,*

$$K = \sum_{n\in I} \lambda_n P_n,$$

where the sum converges in norm.

Proof. Using the functional calculus, we have an isometric $*$-isomorphism $\Psi : C\bigl(\sigma(K)\bigr) \to B(\mathcal{H})$ taking the identity function f to K. Since each λ_n is isolated in $\sigma(K)$ by Lemma 4.15, the characteristic function f_n of $\{ \lambda_n \}$ is continuous on $\sigma(K)$. If $\sigma(K)$ is infinite, then $\lambda_n \to 0$. In any case,

$$f = \sum_{n\in I} \lambda_n f_n$$

in the $\| \cdot \|_\infty$-norm. Therefore, $K = \sum_{n\in I} \lambda_n \Psi(f_n)$. It only remains to show that $\Psi(f_n) = P_n$.

Notice that each $\Psi(f_n)$ is an projection, and $\Psi(f_n)\Psi(f_m) = \Psi(f_n f_m) = 0$ if $n \neq m$. Consequently, $\mathcal{H} = \mathcal{H}_0 \oplus_{n\in I} \mathcal{H}_n$, where $\mathcal{H}_n$ is the range of $\Psi(f_n)$. One can now see that $\Psi(f_n) = P_n$. □

Example 4.17. Let $\mathcal{H} = \ell^2$, and let $\{ e_1, e_2, \dots \}$ be the usual orthonormal basis. Let $\{ \lambda_n \}$ be any sequence tending to 0, and let P_n be the projection onto the space spanned by e_n. Then

$$K = \sum_{n=1}^{\infty} \lambda_n P_n$$

defines a compact operator[17] with spectrum $\{ \lambda_n \}_{n=1}^{\infty} \cup \{ 0 \}$. Notice that if $0 \notin \{ \lambda_n \}$, then 0 is not an eigenvalue of K.

5. Spectral Integrals

Definition 5.1. Let $\mathcal{H}$ be a complex Hilbert space, and suppose $(X, \mathcal{M})$ is a measurable space. Then a $\mathcal{H}$-projection valued measure on $(X, \mathcal{M})$ is a function P from $\mathcal{M}$ into the orthogonal projections in $B(\mathcal{H})$ so that

(a) $P(X) = I$, and
(b) if $\{ E_n \}_{n=1}^{\infty} \subseteq \mathcal{M}$ are pairwise disjoint, then

$$P\Bigl(\bigcup_{n=1}^{\infty} E_n\Bigr) = \sum_{n=1}^{\infty} P(E_n)$$

[16] By a projection, I always mean a self-adjoint idempotent—in other words, an orthogonal projection.

[17] Using the unproved fact that the norm limit of finite rank operators is compact.

in the strong operator topology.[18]

Example 5.2. Let $(X, \mathcal{M}, \mu)$ be a finite measure space, and set $\mathcal{H} = L^2(X, \mu)$. For each $E \in \mathcal{M}$, define $P(E) = M_{\chi_E}$ (see Example 1.1 on page 2). Now if $\{ E_n \} \subseteq \mathcal{M}$ are pairwise disjoint, and we let $E = \bigcup E_n$, then for each $\xi \in L^2(X, \mu)$, the dominated convergence theorem implies that $\sum_{n=1}^{\infty} \xi \chi_{E_n}$ converges to $\xi \chi_E$ in $L^2(X)$. In other words, $\sum_{n=1}^{\infty} P(E_n)$ converges to $P(E)$ in the strong operator topology.

In many respects, projection valued measures act like honest measures. In particular, for each pair of vectors $\xi, \eta \in \mathcal{H}$, the formula

$$\mu_{\xi,\eta}(E) = (P(E)\xi, \eta)$$

defines a complex measure on $(X, \mathcal{M})$. Furthermore, projection valued measures are monotonic: if $F \subseteq E$, then $P(E) = P(F) + P(E \setminus F)$; in particular, we have $P(F) \leq P(E)$ so that $P(F)P(E) = P(F)$ whenever $F \subseteq E$. More generally, a surprising consequence of being projection valued is that

$$P(A \cap B) = P(A)P(B)$$

for all $A, B \in \mathcal{M}$. To see this, notice that

$$\begin{aligned} P(A \cup B) + P(A \cap B) &= P(A) + P(B) \\ P(A \cup B)P(A) + P(A \cap B)P(A) &= P(A) + P(B)P(A) \\ P(A) + P(A \cap B) &= P(A) + P(B)P(A), \end{aligned}$$

where the last bit follows from $P(A \cap B) \leq P(A) \leq P(A \cup B)$.

Just as for ordinary measures, the set of $E \in \mathcal{M}$ for which $P(E) = 0$ is a σ-algebra, and we may enlarge $\mathcal{M}$ if necessary so that $P(E) = 0$ and $A \subseteq E$ implies that $A \in \mathcal{M}$ (and hence $P(A) = 0$). A function $f : X \to \mathbf{C}$ is called measurable if $f^{-1}(V) \in \mathcal{M}$ for every open set $V \subseteq \mathbf{C}$ and we define

$$\|f\|_\infty = \inf\{ k \in [0, \infty] : P\big(\{ x : |f(x)| \geq k \}\big) = 0 \}.$$

Let $\mathcal{L}^\infty(P)$ be the set of equivalence classes of measurable functions with $\|f\|_\infty < \infty$. As usual, $\mathcal{F} = \operatorname{span}\{ \chi_B : B \in \mathcal{M} \}$ are called simple functions. It is not hard to see that $\mathcal{F}$ is dense in the *Banach $*$-algebra* $(\mathcal{L}^\infty(P), \|\cdot\|_\infty)$. (The involution is $f^* = \bar{f}$.)

Define $I : \mathcal{F} \to B(\mathcal{H})$ by

$$I\Big(\sum_{i=1}^{n} \lambda_i \chi_{B_i}\Big) = \sum_{i=1}^{n} \lambda_i P(B_i).$$

Of course, (5) is only well defined because the function $\sum_{i=1}^{n} \lambda_i \chi_{B_i}$ can be rewritten in the form $\sum_{j=1}^{m} \alpha_j \chi_{E_j}$, where the E_j form a disjoint *refinement* of the B_i. Also,

[18]A net $\{ T_\alpha \}$ converges to T in the strong operator topology if $T_\alpha \xi$ converges to $T\xi$ in $\mathcal{H}$ for every $\xi \in \mathcal{H}$.

in this case, it is easy to see that

$$\Big\|I\Big(\sum_{i=1}^{n}\lambda_i\chi_{B_i}\Big)\Big\| = \Big\|\sum_{j=1}^{m}\alpha_i P(E_i)\Big\| = \max_j |\alpha_i|$$
$$= \Big\|\sum \lambda_i\chi_{B_i}\Big\|_\infty.$$

Therefore I is isometric and extends to a linear isometry $I : \mathcal{L}^\infty(P) \to B(\mathcal{H})$ which is usually denoted by

$$I(f) = \int_X f(x)\, dP(x).$$

Notice that $I(f)$ is the *norm* limit of sums of the form $\sum_{i=1}^{n} \lambda_i P(E_i)$.

The following observations are routine for $f \in \mathcal{F}$ and extend to $\mathcal{L}^\infty(P)$ by continuity.

Proposition 5.3. *Suppose $f, g \in \mathcal{L}^\infty(P)$.*

(a) *$\int f\, dP = \int g\, dP$ if and only if $f = g$ almost everywhere.*

(b) *$f \mapsto \int f\, dP$ is an isometric $*$-homomorphism into $B(\mathcal{H})$. That is*

$$(a)\quad \int (f + \lambda g)\, dP = \int f\, dP + \lambda \int g\, dP$$
$$(b)\quad \int fg\, dP = \Big(\int f\, dP\Big)\Big(\int g\, dP\Big)$$
$$(c)\quad \int \bar f\, dP = \Big(\int f\, dP\Big)^*$$
$$(d)\quad \Big\|\int f\, dP\Big\| = \|f\|_\infty.$$

(c) *If $\mu_{\xi,\eta}$ is the complex measure on $(X, \mathcal{M})$ defined by $\mu_{\xi,\eta}(E) = (P(E)\xi, \eta)$, then*

$$\Big(\int f\, dP\xi, \eta\Big) = \int f\, d\mu_{\xi,\eta}, \text{ and}$$
$$\Big\|\Big(\int f\, dP\Big)\xi\Big\|^2 = \int |f|^2\, d\mu_{\xi,\xi}.$$

(d) *$P(A)$ commutes with $(\int f\, dP)$ for all $A \in \mathcal{M}$.*

(e) *$\int f\, dP$ is normal, and is self-adjoint if and only if f is real almost everywhere.*

When dealing with locally compact X, one usually works with regular projection valued measures.

Definition 5.4. A $\mathcal{H}$-projection valued Borel measure P on a locally compact space X is called regular if

$$P(E) = \sup\{\, P(C) : C \text{ is a compact subset of } X\,\}.$$

Remark 5.5. It is immediate that P will be regular if and only if the measures $\mu_{\xi,\xi}$ defined in (5) are regular for all $\xi \in \mathcal{H}$. For our purposes, this is just a technicality as every finite Borel measure on a second countable locally compact space is regular (see for example, Rudin's *Real & Complex Analysis*, 2.18).

Theorem 5.6. *Suppose that A is a unital commutative C^*-subalgebra of $B(\mathcal{H})$, and that Δ is the maximal ideal space of A.*

(a) *There is a unique regular $\mathcal{H}$-projection valued measure P on Δ such that*

$$T = \int_\Delta \widehat{T}\, dP$$

for every $T \in A$ ($\widehat{T}$ is the Gelfand transform of T).

(b) *If $\mathcal{O}$ is open and nonempty in Δ, then $P(\mathcal{O}) \neq 0$.*

(c) *An operator in $S \in B(\mathcal{H})$ commutes with every $T \in A$ if and only if S commutes with $P(E)$ for every Borel set $E \subseteq \Delta$.*

Proof. Uniqueness follows from the fact that P is uniquely determined by the measures $\mu_{\xi,\eta}$, and since the $\mu_{\xi,\eta}$ are regular, the formula

$$\int_\Delta \widehat{T}\, d\mu_{\xi,\eta} = (T\xi, \eta)$$

uniquely determines the $\mu_{\xi,\eta}$ in view of the unicity in the Riesz Representation Theorem.

By Theorem 4.5 on page 10, the inverse of the Gelfand map $M : C(\Delta) \to A$ is an (isometric) $*$-isomorphism into $B(\mathcal{H})$. Let $\mathcal{B}(X)$ denote the bounded Borel functions on X. We want to extend M to a $*$-homomorphism of $\mathcal{B}(X)$ into $B(\mathcal{H})$.

However, given $\xi, \eta \in \mathcal{H}$, the Riesz representation theorem gives us a *unique* regular Borel measure $\mu_{\xi,\eta}$ on Δ so that

$$\big(M(f)\xi, \eta\big) = \int_\Delta f d\mu_{\xi,\eta}$$

for all $f \in C(\Delta)$. Define $M(g)$ for $g \in \mathcal{B}(\Delta)$ by the same formula. It is not hard to see that $M(g)$ is a bounded linear operator.[19] Since $\bar{\mu}_{\xi,\eta} = \mu_{\eta,\xi}$, we have

$$\begin{aligned} \big(\xi, M(f)\eta\big) = \overline{\big(M(f)\eta, \xi\big)} &= \int_\Delta \bar{f}\, d\bar{\mu}_{\eta,\xi} \\ &= \int_\Delta \bar{f}\, d\mu_{\xi,\eta} \\ &= \big(M(\bar{f})\xi, \eta\big) \end{aligned}$$

for all $f \in \mathcal{B}(\Delta)$. Thus $M(f)^* = M(\bar{f})$.

Now we want to see that $M(fg) = M(f)M(g)$ for $f, g \in \mathcal{B}(\Delta)$. But we know this holds for $f, g \in C(\Delta)$, so

$$\int_\Delta fg\, d\mu_{\xi,\eta} = \big(M(f)M(g)\xi, \eta\big) = \int_\Delta f\, d\mu_{M(g)\xi,\eta}$$

for all $f, g \in C(\Delta)$. In particular,

$$g\, d\mu_{\xi,\eta} = d\mu_{M(g)\xi,\eta}$$

[19] Notice that one needs the fact that the $\mu_{\xi,\eta}$ are uniquely determined here. For example, we must have $\mu_{\xi+\zeta,\eta} = \mu_{\xi,\eta} + \mu_{\zeta,\eta}$; it follows that $M(g)(\xi + \zeta) = M(g)\xi + M(g)\zeta$.

for all $\xi, \eta \in \mathcal{H}$ and $g \in C(\Delta)$. Thus, (5) remains valid if $f \in \mathcal{B}(\Delta)$. For convenience, let $\zeta = M(f)^*\eta$. Thus,

$$\begin{aligned}\int_\Delta fg\,d\mu_{\xi,\eta} &= \int_\Delta f\,d\mu_{M(g)\xi,\eta} \\ &= \bigl(M(f)M(g)\xi,\eta\bigr) \\ &= \bigl(M(g)\xi,\zeta\bigr) = \int_\Delta g\,d\mu_{\xi,\zeta}.\end{aligned}$$

Again, the uniqueness of the $u_{\xi,\eta}$'s implies that

$$f\,d\mu_{\xi,\eta} = d\mu_{\xi,\zeta}.$$

Thus $\int_\Delta fg\,d\mu_{\xi,\eta} = \int_\Delta g\,d\mu_{\xi,\zeta}$ for all $f, g \in \mathcal{B}(\Delta)$! Or more simply:

$$\int_\Delta fg\,d\mu_{\xi,\eta} = \bigl(M(f)M(g)\xi,\eta\bigr)$$

for all $f, g \in \mathcal{B}(\Delta)$ and $\xi, \eta \in \mathcal{H}$.

Now suppose $f_n \to f$ pointwise on Δ and that $\{\,\|f_n\|_\infty\,\}$ is bounded. Then, as the $|\mu_{\xi,\eta}|$ are finite measures, it follows that $f_n \to f$ in $L^1(\mu_{\xi,\eta})$ for all $\xi, \eta \in \mathcal{H}$. The point being that I claim $M(f_n) \to M(f)$ in the strong operator topology on Δ. But this follows easily from the fact that

$$\bigl(M(f_n) - M(f)\bigr)^*\bigl(M(f_n) - M(f)\bigr) \to 0$$

in the weak operator topology[20]. To see this, notice that the left hand side of (5) becomes $M(|f_n - f|^2)$, and $|f_n - f|^2 \to 0$ in $L^1(\mu_{\xi,\eta})$ for all $\xi, \eta \in (\mathcal{H})$.

Now if $E \subseteq \Delta$ is Borel, define

$$P(E) = M(\chi_E).$$

By the above, $P(E)$ is a self-adjoint idempotent—i.e., a projection. In particular, $P(X) = M(1) = I$. If $\{\,E_i\,\}_{i=1}^\infty$ are pairwise disjoint with union E, then $f_n = \sum_{i=1}^n \chi_{E_i}$ converges pointwise to $f = \chi_E$, and $\|f_n\|_\infty = 1$ for all n. Thus,

$$P(E) = \sum_{i=1}^\infty P(E_i)$$

in the strong operator topology. In otherwords, P is a $\mathcal{H}$-projection valued measure. Furthermore,

$$\int_\Delta f\,dP = M(f)$$

for all $f \in \mathcal{B}(\Delta)$ as the result is immediate for simple functions (recall that every $f \in \mathcal{B}(\Delta)$ is the uniform limit of uniformly bounded simple functions). This proves (1).

If $\mathcal{O}$ is open and $P(\mathcal{O}) = 0$, then part (1) implies that $T = 0$ if $\widehat{T}$ vanishes off $\mathcal{O}$. Thus, if $P(\mathcal{O}) = 0$, then it follows from Urysohn's lemma that $\mathcal{O} = \emptyset$. This proves (2).

To establish (3), let $\zeta = S^*\eta$. Now by (1), $TS = ST$ for every $T \in A$ if and only if $\mu_{\xi,\zeta} = \mu_{S\xi,\eta}$ for all $\xi, \eta \in \mathcal{H}$. However, this happens if and only if $P(E)S = SP(E)$ for every Borel set E. □

[20] A net $\{\,T_\alpha\,\}$ converges to T in the weak operator topology if $(T_\alpha\xi,\eta) \to (T\xi,\eta)$ for all $\xi, \eta \in \mathcal{H}$.

The following version of the spectral theorem for bounded normal operators now follows quickly.

Corollary 5.7. *If $T \in B(\mathcal{H})$ and $T^*T = TT^*$, then there is a unique $\mathcal{H}$-projection valued measure on the Borel sets of $\sigma(T)$ so that*

$$T = \int_{\sigma(T)} \lambda \, dP(\lambda).$$

In fact, there is an isometric $*$-isomorphism $M : \mathcal{L}^\infty(P) \to B(\mathcal{H})$ such that

$$M(f) = \int_{\sigma(T)} f(\lambda) \, dP(\lambda).$$

An operator $S \in B(\mathcal{H})$ commutes with T if and only if S commutes with every spectral projection $P(E)$.

Remark 5.8. If $S \subseteq B(\mathcal{H})$, then we define

$$S' = \{\, T \in B(\mathcal{H}) : \text{TR=RT for all } R \in S \,\}.$$

With our hypotheses and notation from Theorem 5.6 on page 17, one can show that $M\big(\mathcal{B}(\Delta)\big) = M\big(\mathcal{L}^\infty(P)\big) = A''$. A famous theorem of von-Neumann's (the "Double Commutant Theorem") says that $M\big(C(\Delta)\big)''$ is the strong (or weak) operator closure of $A = M\big(C(\Delta)\big)$.

Furthermore, if $\mathcal{H}$ is separable, then writing $\mathcal{H}$ as a countable direct sum of cyclic subspaces for $A = M\big(C(\Delta)\big)'$ we obtain a cyclic vector for A, and hence a separating vector ξ for $A = M\big(C(\Delta)\big)$ (i.e., $T = 0$ if and only if $T\xi = 0$). The space $L^\infty(\mu_{\xi,\xi})$ coincides with $\mathcal{L}^\infty(P)$.

6. The Holomorphic Symbolic Calculus

The functional calculus for normal elements in a C^*-algebra—both the continuous (Example 4.9 on page 12) and Borel (Theorem 5.6 on page 17)—is quite powerful. Still it is reasonable to ask what can be done with non-normal elements belonging to algebras which may not be C^*-algebras (horrors!). The answer, which comes fairly directly from [8, §§10.21–10.33], will turn out to be "quite a bit," but at the expense of considering a considerably smaller class of functions—namely, holomorphic functions. And except for a comment at the end, we will consider only *unital* algebras. For example, we can define $\exp(x)$ for any x in any unital Banach algebra simply as the sum of the absolutely convergent series $\sum_{n=0}^\infty x^n/n!$. Naturally, we can do a bit more than that. Furthermore, these techniques will yield non-trivial results even for $n \times n$ matrices over $\mathbf{C}$. For example, it will follow from Theorem 6.11 on page 27 that *any* invertible $n \times n$ matrix has a logarithm. First, it will be convenient to digress slightly and make a few comments about Banach space valued integrals of a rather special sort.

Suppose that A is a Banach space and that $f : [a, b] \to A$ is continuous. Then one can define

$$\int_a^b f(t) \, dt$$

is a number of ways.[21] I will describe a very straightforward way to do so here.

[21] See [11, Lemma 1.91 and footnote 21].

Just as in the scalar case, if $\mathcal{P} = \{\, a = t_0 < t_1 < \cdots < t_n = b \,\}$ is a *partition* of $[a,b]$, then $\|\mathcal{P}\| = \max_{1\le i\le n} \Delta t_i = \max_{1\le i\le n} |t_i - t_{i-1}|$ is called the *mesh* of $\mathcal{P}$. One says that $\mathcal{P}'$ is a *refinement* of $\mathcal{P}$ and writes $\mathcal{P}' \prec \mathcal{P}$ if $\mathcal{P}' \supseteq \mathcal{P}$. If $\zeta = (z_1, z_2, \ldots, z_n) \in [a,b]^n$ satisfies $z_i \in [t_{i-1}, t_i]$, then

$$\mathcal{R}(f, \mathcal{P}, \zeta) = \sum_{i=1}^{n} f(z_i)\Delta t_i$$

is called a *Riemann sum* for f with respect to $\mathcal{P}$. Since f is necessarily uniformly continuous, notice that given an $\epsilon > 0$ there is a $\delta > 0$ so that if $\|\mathcal{P}\| < \delta$ and $\mathcal{P}' \prec \mathcal{P}$, then

$$\|\mathcal{R}(f, \mathcal{P}', \zeta') - \mathcal{R}(f, \mathcal{P}, \zeta)\| < \epsilon$$

for any appropriate ζ and ζ'[22].

Now for each $n \in \mathbf{N}$, let $\mathcal{P}_n$ be the uniform partition of $[a,b]$ into 2^n sub-intervals. Let $\zeta_n = (t_0, t_1, \ldots, t_{2^n - 1})$, and put

$$a_n = \mathcal{R}(f, \mathcal{P}_n, \zeta_n).$$

Notice that $m \ge n$ implies that $\mathcal{P}_m \prec \mathcal{P}_n$. It now follows from (6) that $\{\, a_n \,\}_{n=1}^{\infty}$ is Cauchy in A. We will define

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} a_n.$$

Using (6) again, it is a simple matter to check that for all $\epsilon > 0$ there is a $\delta > 0$ so that $\|\mathcal{P}\| < \delta$ implies that

$$\|\mathcal{R}(f, \mathcal{P}, \zeta) - \int_a^b f(t)\,dt\| < \epsilon$$

for *any* compatible vector ζ.

These observations ought to suffice for most purposes. For example, it is routine to verify that if A is a Banach algebra, if $f : [a,b] \to A$ is continuous, and if $x \in A$, then

$$x \int_a^b f(t)\,dt = \int_a^b x f(t)\,dt, \text{ and}$$
$$\Bigl(\int_a^b f(t)\,dt\Bigr) x = \int_a^b f(t) x\,dt.$$

Our interest in these sorts of integrals is to define contour integrals

$$\int_\Gamma f(\lambda)\,d\lambda,$$

[22]To see this, simply choose δ so that $|x - y| < \delta$ implies that $\|f(x) - f(y)\| < \epsilon/(b-a)$. Then the left-hand side of (6) is less than or equal to $\sum_{i=1}^{n} \|\mathcal{R}(f, \mathcal{P}_i', \zeta^i) - f(z_i)\Delta t_i\|$, where $\mathcal{P}_i' = \mathcal{P}' \cap [t_{i-1}, t_i]$. If $\mathcal{P}_i' = \{\, t_{i-1} = t_0^i < \cdots < t_{m_i}^i = t_i \,\}$, then (6) is bounded by

$$\begin{aligned}\sum_{i=1}^{n} \Bigl\| \sum_{k=1}^{m_i} f(z_k^i)\Delta t_k^i - f(z_i) \sum_{k=1}^{m_i} \Delta t_k^i \Bigr\| &\le \sum_{i=1}^{n}\sum_{k=1}^{m_i} \|f(z_k^i) - f(z_i)\|\Delta t_k^i \\ &\le \epsilon/(b-a) \sum_{i=1}^{n}\sum_{k=1}^{m_i} \Delta t_k^i = \epsilon.\end{aligned}$$

where Γ is a piecewise smooth path in $\mathbf{C}$ (not necessarily connected), and f is a continuous function from (the image of) Γ into a Banach algebra A. Observe that it is clear from our definition of the integral that if $\Lambda \in A^*$, then

$$\Lambda\Big(\int_\Gamma f(\lambda)\,d\lambda\Big) = \int_\Gamma \Lambda\big(f(\lambda)\big)\,d\lambda.$$

Therefore the usual Cauchy Theorem [9, Theorem 10.35] applies virtually word-for-word to A-valued holomorphic functions. We will also make use of the following topological fact. If K is a compact subset of $\mathbf{C}$ and if Ω is an open neighborhood of K, then there is a contour Γ which *surrounds* K *in* Ω in the sense that Γ lies in Ω and

$$\operatorname{ind}_\Gamma(z) := \frac{1}{2\pi i}\int_\Gamma \frac{1}{\xi - z}\,d\xi = \begin{cases} 1 & \text{if } z \in K, \text{ and} \\ 0 & \text{if } z \notin \Omega. \end{cases}$$

This follows from the proof of [9, Theorem 13.5]. Notice that Γ may have to have several components.[23]

The basic idea in the following will be to use such contours as described in the previous paragraph with $K = \sigma(x)$, the spectrum of an element x in a Banach algebra, to make sense of a kind of A-valued Cauchy Integral formula. The first step is to see that we get the "right" thing for functions of the form $f(\lambda) = (\alpha - \lambda)^n$.

Lemma 6.1. *Suppose that A is a unital Banach algebra, that $x \in A$, and that $\alpha \in \mathbf{C} \setminus \sigma(x)$. Then, if Γ is any contour surrounding $\sigma(x)$ in the complement of α in $\mathbf{C}$, and if n is any integer,*

$$\frac{1}{2\pi i}\int_\Gamma (\alpha - \lambda)^n[\lambda e - x]^{-1}\,d\lambda = [\alpha e - x]^n. \tag{6.1}$$

Remark 6.2. Notice that $\lambda \mapsto [\lambda e - x]^{-1}$ is holomorphic (hence continuous) off $\sigma(x)$ so that the integral is defined. Similarly, the right-hand side makes sense for every $n \in \mathbf{Z}$ since $\alpha \notin \sigma(x)$.

Proof. Let the value of (6.1) be y_n. If $\lambda \notin \sigma(x)$, then

$$\begin{aligned}[]
[\lambda e - x]^{-1} - [\alpha e - x]^{-1} &= [\lambda e - x]^{-1}\big((\alpha e - x) - (\lambda e - x)\big)[\alpha e - x]^{-1} \\
&= (\alpha - \lambda)[\lambda e - x]^{-1}[\alpha e - x]^{-1} \\
&= (\alpha - \lambda)[\alpha e - x]^{-1}[\lambda e - x]^{-1}.
\end{aligned}$$

Thus

$$y_n = \frac{[\alpha e - x]^{-1}}{2\pi i}\left(\int_\Gamma (\alpha - \lambda)^n\,d\lambda + \int_\Gamma (\alpha - \lambda)^{n+1}[\lambda e - x]^{-1}\,d\lambda\right).$$

The first integral is always zero when $n \neq -1$, and equals zero even when $n = -1$ because we have assumed that $\operatorname{ind}_\Gamma(\alpha) = 0$. Thus,

$$(\alpha e - x)y_n = y_{n+1}. \tag{6.2}$$

Thus (6.1) will follow from (6.2) once we show that

$$\frac{1}{2\pi i}\int_\Gamma [\lambda e - x]^{-1}\,d\lambda = e.$$

[23]This definition is more subtle that it seems at first glace. Notice that if K is the unit circle and Ω is the complement of 0 in $\mathbf{C}$, then Γ will have to have two components; for example, Γ could consist of a positively oriented circle of radius $3/2$ and a negatively oriented circle of radius $1/2$.

Notice that if Γ_r is a positively oriented circle of radius $r > \|x\|$, then Γ_r surrounds $\sigma(x)$ and

$$[\lambda e - x]^{-1} = \sum_{n=0}^{\infty} \lambda^{-n-1} x^n$$

uniformly for $\lambda \in \Gamma_r$. Thus

$$\frac{1}{2\pi i} \int_{\Gamma_r} [\lambda e - x]^{-1} \, d\lambda = e \tag{6.3}$$

by termwise integration. Finally, since $f(\lambda) = [\lambda e - x]^{-1}$ is holomorphic on the complement of $\sigma(x)$, and since

$$\mathrm{ind}_{\Gamma}(z) = 1 = \mathrm{ind}_{\Gamma_r}(z)$$

for all $z \in \sigma(x)$, Cauchy's Theorem[24] implies that (6.3) holds with Γ_r replaced by Γ. $\square$

Now suppose that R is a rational function with poles at $\alpha_1, \alpha_2, \dots, \alpha_n$. The usual theory of partial fraction decomposition implies that

$$R(\lambda) = P(\lambda) + \sum_{k,m} c_{m,k} (\lambda - \alpha_m)^{-k}, \tag{6.4}$$

where P is a polynomial and each $c_{m,k}$ is a complex constant. Note that the sum in (6.4) is finite. If A is a unital Banach algebra and if $x \in A$ has spectrum disjoint from the poles of R, then for the sake of definiteness we define

$$R(x) = P(x) + \sum_{k,m} c_{m,k} [x - \alpha_m e]^{-k}.$$

Observe that if α and α' are not in $\sigma(x)$, then $[\alpha e - x]^{-1}$ commutes with $[\alpha' e - x]^{-1}$ as well as with $Q(x)$ for any polynomial Q. Thus if (6.4) also equals

$$R(\lambda) = \frac{P_1(\lambda)}{(\lambda - \alpha_1)^{r_1} \cdots (\lambda - \alpha_n)^{r_n}},$$

then $R(x) = P_1(x)[x - \alpha_1 e]^{-r_1} \cdots [x - \alpha_n e]^{-r_n}$. Thus, the definition of $R(x)$ is independent of the representation of R. In particular, if $R(\lambda) = f(\lambda) g(\lambda)$ with f and g rational, then $R(x) = f(x) g(x)$. We will need this observation for our main result (Theorem 6.5 on the following page). However combining Lemma 6.1 on the preceding page and (6.4), we obtain the next result.

Theorem 6.3. *Suppose that A is a unital Banach algebra and that $x \in A$. Let Ω be a neighborhood of $\sigma(x)$ in $\mathbf{C}$, and let Γ be a contour surrounding $\sigma(x)$ in Ω. Then*

$$R(x) = \frac{1}{2\pi i} \int_{\Gamma} R(\lambda) [\lambda e - x]^{-1} \, d\lambda$$

for all rational functions $R \in H(\Omega)$.

Since Theorem 6.3 implies that the Cauchy formula gives the "right" answer for rational functions, we are lead to the next definition.

---

[24] We need a fairly sophisticated version of Cauchy's Theorem here — for example, [9, Theorem 10.35] will do.

Definition 6.4. Suppose that A is a unital Banach algebra and that $\Omega \subseteq \mathbf{C}$ is open. Let

$$A_\Omega = \{ x \in A : \sigma(x) \subseteq \Omega \}.$$

Also if $f \in H(\Omega)$ and $x \in A_\Omega$, then we define

$$\tilde{f}(x) = \frac{1}{2\pi i} \int_\Gamma f(\lambda)[\lambda e - x]^{-1} \, d\lambda$$

for any contour Γ which surrounds $\sigma(x)$ in Ω. The collection of all such functions $\tilde{f} : A_\Omega \to A$ will be denoted by $\tilde{H}(A_\Omega)$.

Some comments on this definition are in order. First $\tilde{f}(x) \in A$ since the integrand is continuous and A is complete. Secondly, since the integrand is actually holomorphic in $\Omega \setminus \sigma(x)$, Cauchy's theorem implies that $\tilde{f}(x)$ is independent of the choice of contour Γ (provided Γ surrounds $\sigma(x)$). A consequence of this that we will make use of without comment, is that $\tilde{f}(x)$ *depends only on the germ of f restricted to $\sigma(x)$*. Finally, if R is a rational function in $H(\Omega)$ and if $x \in A_\Omega$, then $\tilde{R}(x) = R(x)$ by Theorem 6.3 on the previous page.

This brings us to the main result.

Theorem 6.5. *Let A be a unital Banach algebra and let $\Omega \subseteq \mathbf{C}$ be open. Then $\tilde{H}(A_\Omega)$ is a complex algebra, and $f \mapsto \tilde{f}$ is an algebra isomorphism of $H(\Omega)$ onto $\tilde{H}(A_\Omega)$. Furthermore, if f_n converges to f uniformly on compact subsets of Ω in $H(\Omega)$, then*

$$\tilde{f}(x) = \lim_{n \to \infty} \tilde{f}_n(x)$$

for all $x \in A_\Omega$.

Proof. Clearly $f \mapsto \tilde{f}$ is linear, and it is onto by definition. Furthermore if $\tilde{f}$ is the zero function, then

$$\begin{aligned} \tilde{f}(\alpha e) &= \frac{1}{2\pi i} \int_\Gamma f(\lambda)(\lambda e - \alpha e)^{-1} \, d\lambda \\ &= \frac{1}{2\pi i} \int_\Gamma \frac{f(\lambda)}{\lambda - \alpha} \, d\lambda \cdot e \\ &= f(\alpha) e. \end{aligned}$$

It follows that $f = 0$. Therefore $f \mapsto \tilde{f}$ is one-to-one. Since $\lambda \mapsto \|[\lambda e - x]^{-1}\|$ is bounded on any appropriate contour Γ, the continuity assertion is a straightforward consequence of the definition.

For the remaining statements, it will suffice to show that $f \mapsto \tilde{f}$ is multiplicative. So suppose that $h \in H(\Omega)$ is of the form $h = fg$ with $f, g \in H(\Omega)$. If f and g are rational functions, then

$$\tilde{h}(x) = \tilde{f}(x)\tilde{g}(x)$$

in view of the comments preceding Theorem 6.3 on the preceding page. But Runge's Theorem [8, Theorem 13.9][25] implies that there are rational functions $\{ f_n \}$ and $\{ g_n \}$ in $H(\Omega)$ so that $f_n \to f$ and $g_n \to g$ uniformly on compact subsets of Ω.

[25]**Runge's Theorem**: Let Ω be an open set in the plane, let A be a set which contains one element of each component of $S^2 \setminus \Omega$, and assume that $f \in H(\Omega)$. Then there is a sequence of rational functions $\{ r_n \}$, with poles only in A such that $r_n \to f$ uniformly on compacta in Ω.

It should be remarked here that in the special case where Ω is simply connected (and therefore $S^2 \setminus \Omega$ is connected by [9, Theorem 13.11]), we may take $A = \{ \infty \}$. Then each r_n is a polynomial.

Clearly $h_n = f_n g_n$ converges to h uniformly on compacta as well, and the result follows from the continuity proved above. $\square$

Remark 6.6. Suppose that $x \in A_\Omega$ and that $f \in H(\Omega)$ is given by a convergent power series in Ω. Say $f(z) = \lim_{n\to\infty} p_n(z)$ for all $z \in \Omega$, where $p_n(z) = \sum_{k=0}^{n} a_n(z - z_0)^n$. Then Theorem 6.5 on the preceding page implies that $p_n(x)$ converges to $\tilde{f}(x)$. (This follows even if $\|x - z_0 e\|$ is not inside the circle of convergence of $\sum_{k=0}^{\infty} a_n z^n$!) Thus there is no ambiguity when considering expressions like $\exp(x)$. (If you like, $\sum_{n=0}^{\infty} x^n/n! = \widetilde{\exp}(x)$.)

Theorem 6.7. *Let A be a unital Banach algebra, Ω open in $\mathbf{C}$, $x \in A_\Omega$, and $f \in H(\Omega)$. Then*

(a) *$\tilde{f}(x)$ is invertible in A if and only if $f(\lambda) \neq 0$ for all $\lambda \in \sigma(x)$.*
(b) *$\sigma\big(\tilde{f}(x)\big) = f\big(\sigma(x)\big)$.*

Proof. If $f(\lambda) \neq 0$ for all $\lambda \in \sigma(x)$, then $g = 1/f \in H(\Omega_1)$ for some open set satisfying $\sigma(x) \subseteq \Omega_1 \subseteq \Omega$. Now $fg = 1$ in $H(\Omega_1)$, so $\tilde{f}(x)\tilde{g}(x) = e$ in A_{Ω_1}. That is, $\tilde{f}(x)$ is invertible.

Conversely, if $f(\alpha) = 0$ for some $\alpha \in \sigma(x)$, then there is a $h \in H(\Omega)$ such that $(\lambda - \alpha)h(\lambda) = f(\lambda)$. Thus

$$(x - \alpha e)\tilde{h}(x) = \tilde{f}(x) = \tilde{h}(x)(x - \alpha e).$$

Since $(x - \alpha e)$ is not invertible, then neither is $\tilde{f}(x)$. This proves part (a).

For part (b), fix $\beta \in \mathbf{C}$. Note that $\beta \in \sigma\big(\tilde{f}(x)\big)$ if and only if $\tilde{f}(x) - \beta e$ is not invertible. By part (a), this is the case if and only if the function $f - \beta$ has a zero in $\sigma(x)$. That is, if and only if $\beta \in f\big(\sigma(x)\big)$. $\square$

The second part of the previous theorem is called *the Spectral Mapping Theorem*. It will play an important rôle in the next result which shows that the holomorphic calculus behaves nicely with respect to composition.

Theorem 6.8. *Suppose that A is a unital Banach algebra, that Ω is open in $\mathbf{C}$, that $x \in A_\Omega$, and that $f \in H(\Omega)$. Also suppose that Ω_1 is an open neighborhood of $\sigma\big(\tilde{f}(x)\big)$, and that $g \in H(\Omega_1)$. Finally let $\Omega_0 = \{\,\lambda \in \Omega : f(\lambda) \in \Omega_1\,\}$, and define $h \in H(\Omega_0)$ by $h(\lambda) = g\big(f(\lambda)\big)$. Then $\tilde{f}(x) \in A_{\Omega_1}$, and $\tilde{h}(x) = \tilde{g}\big(\tilde{f}(x)\big)$.*

Remark 6.9. More simply put, if $h = g\circ f$, then $\tilde{h} = \tilde{g}\circ\tilde{f}$. It should also be observed that, although this result is reasonably straightforward for rational functions, we have only shown that $\tilde{f}_n \to \tilde{f}$ pointwise. Thus the proof can not proceed by rational approximation.

Proof. The Spectral Mapping Theorem implies that $\sigma\big(\tilde{f}(x)\big) \subseteq \Omega_1$. Therefore $\tilde{f}(x) \in A_{\Omega_1}$ and $\tilde{g}\big(\tilde{f}(x)\big)$ is at least defined.

Let Γ_1 be a contour which surrounds $f\big(\sigma(x)\big)$ in Ω_1. Since $z \mapsto \operatorname{ind}_{\Gamma_1}(z)$ is continuous, there is an open set W such that $\sigma(x) \subseteq W \subseteq \Omega_0$ with $\operatorname{ind}_{\Gamma_1}\big(f(\lambda)\big) = 1$ for all $\lambda \in W$. (In particular, $\lambda \in W$ implies that $f(\lambda) \notin \Gamma_1$.) Now let Γ_0 be

---

On the other hand, by considering $f(z) = 1/z$ in $\Omega = \mathbf{C}\setminus\{\,0\,\}$, we see that it is not usually possible to approximate functions uniformly on compacta by polynomials if Ω is not simply connected. (In fact, such approximations are possible *if and only if* Ω is simply connected [9, Theorem 13.11].)

a contour in W which surrounds $\sigma(x)$. For each $\zeta \in \Gamma_1$, define $\varphi_\zeta \in H(W)$ by $\varphi_\zeta(\lambda) = 1/\bigl(\zeta - f(\lambda)\bigr)$. Thus

$$[\zeta e - \tilde{f}(x)]^{-1} = \tilde{\varphi}_\zeta(x) = \frac{1}{2\pi i}\int_{\Gamma_0} \bigl(\zeta - f(\lambda)\bigr)^{-1}[\lambda e - x]^{-1}\,d\lambda$$

for all $\zeta \in \Gamma_1$. But

$$\begin{aligned}
\tilde{g}\bigl(\tilde{f}(x)\bigr) &= \frac{1}{2\pi i}\int_{\Gamma_1} g(\zeta)[\zeta e - \tilde{f}(x)]^{-1}\,d\zeta \\
&= \frac{1}{2\pi i}\int_{\Gamma_1} \frac{g(\zeta)}{2\pi i}\int_{\Gamma_0} \bigl(\zeta - f(\lambda)\bigr)^{-1}[\lambda e - x]^{-1}\,d\lambda\,d\zeta \\
&= \frac{1}{2\pi i}\int_{\Gamma_0} \left(\frac{1}{2\pi i}\int_{\Gamma_1} g(\zeta)\bigl(\zeta - f(\lambda)\bigr)^{-1}\,d\zeta\right)[\lambda e - x]^{-1}\,d\lambda,
\end{aligned}$$

which, by the Cauchy integral formula, is

$$\begin{aligned}
&= \frac{1}{2\pi i}\int_{\Gamma_0} g\bigl(f(\lambda)\bigr)[\lambda e - x]^{-1}\,d\lambda \\
&= \tilde{h}(x).
\end{aligned}$$

□

Remark 6.10 (Separating 0 from ∞). In the next result, we require the hypothesis that "$\sigma(x)$ does not separate 0 from ∞". What does this mean? (This has caused me some annoyance!) Technically, it means that 0 is in the unique unbounded connected component of $\mathbf{C} \setminus \sigma(x)$. (Since $\sigma(x)$ is compact and therefore bounded and since $\{\, z \in C : |z| > r \,\}$ is connected, the complement of $\sigma(x)$ has a unique unbounded connected component.) For the proof below, I would like to believe that when $\sigma(x)$ separates 0 and ∞ we can conclude that there is a simply connected *region* Ω containing $\sigma(x)$ and not containing zero. Here "region" is defined to be an open and *connected* subset of the plane. However, proving this seems to me to be pretty hard. We can certainly claim that 0 and ∞ lie in the same component of the complement of $\sigma(x)$ in the Riemann sphere S^2. Thus there is a path γ connecting 0 to ∞ in $S^2 \setminus \sigma(x)$. I would like to claim that $\Omega := S^2 \setminus \gamma^*$ is simply connected by [9, Theorem 13.11].[26] Unfortunately, Rudin's result assumes that Ω is a region and it is easy to see that γ might have loops so that Ω need not be connected. However, such a path γ can only cross itself finitely many times (as a curve on S^2!). Therefore we can "snip" off the loops and assume that γ^* is homeomorphic to the unit interval I. Then it appears to be nontrivial, but none the less true, that $S^2 \setminus \gamma^*$ is connected. This is a consequence of the homology machinery used to prove the Jordan Curve Theorem; for example, see [4, Theorem 63.2], or if your dare, [10, Lemma 4.7.13]. With this result in hand, $\Omega = S^2 \setminus \gamma^*$ is connected, and we can apply Rudin's Theorem 13.11 with a clear conscience. Alternatively, the proof of [9, Theorem 13.11] seems to imply that, even if γ^* is not homeomorphic to I, the connectedness of $S^2 \setminus \Omega$, where $\Omega := S^2 \setminus \gamma^*$ as above, implies that every closed path in Ω is homotopic to a point. (Thus, Ω is a disjoint union of simply connected regions.) Then, since $0 \notin \Omega$, we can find branches of the logarithm on each component and the proof of Theorem 6.11 proceeds just fine. I find neither of these "solutions" entirely satisfactory, but it's the best I can do now.

[26]Here I am using Rudin's formalism, and using γ^* to denote the point set associated to the curve γ.

Theorem 6.11. *Suppose that A is a unital Banach algebra, that $x \in A$ is invertible, and that $\sigma(x)$ does not separate 0 from ∞. Then*

(a) *There is a logarithm for x; that is, there is a $y \in A$ such that $\exp(y) = x$.*
(b) *There are roots of all orders for x; that is given $n \in \mathbf{Z} \setminus \{0\}$, there is a $z \in A$ such that $z^n = x$.*
(c) *For each $\epsilon > 0$ there is a polynomial P such that $\|x^{-1} - P(x)\| < \epsilon$.*

Remark 6.12. This result is non-trivial even for $n \times n$ matrices. For example as mentioned in the beginning of the section, every invertible $n \times n$ complex matrix has a logarithm.

Proof. By hypothesis, 0 lies in the unbounded component of $\mathbf{C} \setminus \sigma(x)$. Thus there is a simply connected set Ω such that $\sigma(x) \subseteq \Omega$ and $0 \notin \Omega$ (see Remark 6.10 above). Therefore, there is a $f \in H(\Omega)$ satisfying

$$\exp\bigl(f(\lambda)\bigr) = \lambda$$

for all $\lambda \in \Omega$. Let $y = \tilde{f}(x)$. Then $\exp(y) = x$ by the preceding result. This proves part (a).

Now if $z = \exp(y/n)$, then $z^n = x$. So part (b) follows from part (a).

Finally since Ω is simply connected, Runge's Theorem implies that $f(\lambda) = 1/\lambda$ can be approximated uniformly on compacta on Ω by polynomials. Now part (c) follows. □

Remark 6.13. In the event A is a non-unital Banach algebra, then we can let A^+ be the usual unital Banach algebra containing A as a co-dimension one ideal. (Recall that if $x \in A$, then $\sigma(x) = \sigma_{A^+}(x)$ by definition.) Then if $\pi : A^+ \to \mathbf{C}$ is the quotient map, it is immediate from the definition of the integral that

$$\pi\Bigl(\int_\Gamma f(\lambda)\,d\lambda\Bigr) = \int_\Gamma \pi\bigl(f(\lambda)\bigr)\,d\lambda.$$

In particular if $x \in A^+_\Omega$ and if Γ surrounds $\sigma(x)$ in Ω, then

$$\pi\bigl(\tilde{f}(x)\bigr) = \frac{1}{2\pi i}\int_\Gamma \frac{f(\lambda)}{\lambda - \pi(x)}\,d\lambda = f\bigl(\pi(x)\bigr).$$

It follows that $\tilde{f}(x) \in A$ if and only if $f\bigl(\pi(x)\bigr) = 0$. Thus if $x \in A_\Omega$, then we can define $\tilde{f}(x)$ for all $f \in H(\Omega)$ which vanish at 0.

Remark 6.14 (Connections with the "ordinary" functional calculus). Suppose now that A is a unital C^*-algebra and that x is a normal element in A. (In these notes, we've only considered $A = B(H)$ and the functional calculus as in Example 4.9.) Suppose that $f \in C\bigl(\sigma(x)\bigr)$ is such that there is an open neighborhood Ω of $\sigma(x)$ and a $h \in H(\Omega)$ such that $h|_{\sigma(x)} = f$. We want to see that $\tilde{h}(x) = f(x)$, where of course, $f(x)$ denotes the element of A produced via the usual functional calculus for C^*-algebras. However, if h is a rational function, then it is clear that $\tilde{h}(x) = h(x) = f(x)$. But there are rational functions h_n such that $h_n \to h$ uniformly on compact subsets of Ω. Therefore $h_n(x) \to \tilde{h}(x)$ in A. On the other hand, if $f_n := h_n|_{\sigma(x)}$, then $f_n \to f$ uniformly on $\sigma(x)$. Therefore $f_n(x) \to f(x)$. Since we agreed that $h_n(x) = f_n(x)$, the assertion follows.

7. A Spectral Theorem for Unbounded Operators

Here we want to look at a version of the spectral theorem for unbounded *self-adjoint* operators in a Hilbert space $\mathcal{H}$. We are only going to skim the surface. I got this approach from a course I took from Marc Rieffel long ago when my hat covered a good deal more hair and a younger man's brain. (Ok, it was in the spring of 1978.) Much of what's in here, and a good deal more, can be found in [8, Chap. 13].

As Rieffel did, we'll start with a discussion that motivates our limited point of view. This section could easily be expanded to include normal operators, but that will be left to the interested reader.

7.1. **Stone's Theorem: Part I.** A homomorphism u of $\mathbf{R}$ into the unitary group $U(\mathcal{H})$ of a Hilbert space $\mathcal{H}$ equipped with the strong operator topology is variously called a *one-parameter unitary group* or a *unitary representation*. We'll use the former terminology since we are only concerned with the group $\mathbf{R}$. We can also think of u as defining a $*$-homomorphism, also denoted by u, of $C_c(\mathbf{R})$ into $B(\mathcal{H})$:

$$u(f) = \int_{\mathbf{R}} f(r)u_r\,dr.$$[27]

The image of $C_c(\mathbf{R})$ under u generates a commutative C^*-subalgebra of $B(\mathcal{H})$ which is isomorphic to $C_0(X)$ for second countable locally compact space X via the Gelfand transform: $T \mapsto \widehat{T}$ or $u(f) \mapsto \hat{u}(f)$. We can decompose $\mathcal{H}$ into countably many cyclic subspaces $\{\,\mathcal{H}_i\,\}_{i\in I}$ with cyclic vector z_i for u. For each i the map

$$T \mapsto (Tz_i \mid z_i)$$

determines a positive linear functional on $C_0(X)$. Hence there is a Radon measure μ_i on X such that

$$\big(u(f)z_i \mid z_i\big) = \int_X \hat{u}(f)(x)\,d\mu_i(x) \quad \text{for all } f \in C_c(\mathbf{R}).$$

Since

$$\big(u(f)z_i \mid u(g)z_i\big) = \big(u(g^* * f)z_i \mid z_i\big) = \int_X \hat{u}(f)(x)\overline{\hat{u}(g)(x)}\,d\mu_i(x),$$

$u(f)z_i \mapsto \hat{u}(f)$ induces a unitary isomorphism of $\mathcal{H}_i$ onto $L^2(X,\mu_i)$ which intertwines the restriction of $u(f)$ to $\mathcal{H}_i$ with multiplication by $\hat{u}(f)$. We can then let $Y = \{\,(x,i) : x \in X \text{ and } i \in I\,\}$ be the disjoint union of suitably many copies of X and let μ be the corresponding Radon measure on Y (induced by the μ_i) so that $H \cong \bigoplus_i L^2(X,\mu_i)$ is isomorphic to $L^2(Y,\mu)$ via a unitary which intertwines $u(f)$ and multiplication $M_{\tilde{f}}$ by a continuous function $\tilde{f} : Y \to \mathbf{C}$ given by $\tilde{f}(x,i) = \hat{u}(f)(x)$.

---

[27]Since u is only strongly continuous, and not norm-continuous, it takes a bit more fussing in order to make sense out of integrals of this type than we discussed, for example, in Section 6. However, if $v \in \mathcal{H}$, then $r \mapsto u_r v$ is continuous from $\mathbf{R}$ into $\mathcal{H}$ with its norm topology, and the integral above is to be interpreted as

$$u(f)v = \int_{\mathbf{R}} f(r)u_r v\,dr.$$

If $y \in Y$, then $f \mapsto \tilde{f}(y)$ is a linear functional on $C_c(\mathbf{R})$ such that

$$|\tilde{f}(y)| \leq \|\tilde{f}\|_\infty = \|u(f)\| \leq \|f\|_1.$$

Therefore the linear functional is bounded, and since, $(f * g)\tilde{}(y) = \tilde{f}\tilde{g}(y) = \tilde{f}(y)\tilde{g}(y)$, it extends to a complex homomorphism on $L^1(\mathbf{R})$. Such things are always given by integration against a character:

$$\tilde{f}(y) = \int_{-\infty}^{\infty} f(r)e^{-ih(y)r}\,dr.$$

Thus we obtain a function $h : Y \to \mathbf{R}$ such that $\tilde{f}(y) = \hat{f}\big(h(y)\big)$ — where $\hat{f}$ is the Fourier transform of f. Since the $\hat{f}$ generate the topology on $\widehat{\mathbf{R}} = \mathbf{R}$, it follows that h is continuous.

Since $u_r u(f) = u\big(\lambda(r)f\big)$, where $\lambda(r)f(s) = f(s-r)$, and since $\big(\lambda(r)f\big)^\wedge(y) = e^{-iry}\hat{f}(y)$, the isomorphism of $\mathcal{H}$ with $L^2(Y,\mu)$ intertwines u_r with $M_{e^{-irh(\cdot)}}$. Therefore we think of u_r as $\exp(-irA)$ where A is the "operator" M_h. The point is that h is not usually a bounded function so that A is not bounded — in fact, A is not even everywhere defined!

Nevertheless, we have sketched a proof of the following.

Theorem 7.1 (Stone's Theorem (first half))**.** *Let $\{\, u_r \,\}_{r\in\mathbf{R}}$ be a one-parameter group of unitaries on $\mathcal{H}$. Then there is a (second countable) locally compact space Y and a Radon measure μ on Y together with a (possibly unbounded) continuous real-valued function h on Y and a unitary of $\mathcal{H}$ onto $L^2(Y,\mu)$ intertwining u_r and $M_{e^{irh(\cdot)}}$.*

Of course, the "operator" $A = M_h$ corresponds to an "operator" of some kind on the original Hilbert space $\mathcal{H}$ and we want to get our hands on these sorts of operators.

Example 7.2. Let $\mathcal{H} = L^2(\mathbf{R})$ and let u be the *left-regular representation*:

$$u_r f(s) = f(s-r).$$

In this case, we can identify Y with $\mathbf{R}$ and $h(r) = r$ for all $r \in \mathbf{R}$. Since $(i\frac{df}{dx})^\wedge(r) = r\hat{f}(r)$, M_h corresponds to the operator $i\frac{d}{dx}$.

7.2. **Unbounded Operators.** Motivated by $T = \frac{d}{dx}$ on $L^2(\mathbf{R})$, we make the following definition.

Definition 7.3. A (possibly unbounded) operator in $\mathcal{H}$ is a linear map $T : \mathcal{D}(T) \subset \mathcal{H} \to \mathcal{H}$, where $\mathcal{D}(T)$ is a subspace of $\mathcal{H}$.

Remark 7.4. We are usually only interested in *densely defined* operators — that is, operators T where $\mathcal{D}(T)$ is dense in $\mathcal{H}$. More specifically, we want to know when such an operator is (unitarily equivalent to) a multiplication operator such that which arose in the previous subsection.

Example 7.5. Let $h \in C(Y)$ be a continuous function on a locally compact Hausdorff space Y, let μ and be a Radon measure on Y and let M_h be the multiplication operator on $L^2(Y,\mu)$. Then the natural choice for $\mathcal{D}(M_h)$ is

$$D(M_h) = \{\, f \in L^2(Y,\mu) : hf \in L^2(Y,\mu) \,\}.$$

Since $C_c(Y) \subset \mathcal{D}(M_h)$, $\mathcal{D}(M_h)$ is dense. Furthermore, if $f, g \in \mathcal{D}(M_h)$, then certainly $g \in \mathcal{D}(M_{\bar{h}})$ and

$$(M_h f \mid g) = (f \mid M_{\bar{h}} g).$$

Suppose that $g \in L^2(Y, \mu)$ is such that

$$f \mapsto (M_h f \mid g)$$

is continuous on $\mathcal{D}(M_h)$. Since the map is *linear* and continuous, it is bounded, and therefore extends to a bounded linear functional on $L^2(Y, \mu)$. Therefore the Riesz Representation Theorem implies that there is a $g' \in L^2(Y, \mu)$ such that

$$(M_h f \mid g) = (f \mid g') \quad \text{for all } f \in \mathcal{D}(M_h).$$

Let Φ be the linear functional on $C_c(Y) \subset L^2(Y, \mu)$ given by integration against the function $M_{\bar{h}} g$. Then

$$\begin{aligned} \Phi(g) &= (f \mid M_{\bar{h}} g) \\ &= (M_h f \mid g) \\ &= (f \mid g') \\ &\leq \|f\|_2 \|g'\|_2. \end{aligned}$$

Thus Φ determines a bounded linear functional on $L^2(Y, \mu)$ that agrees with integration against g' on a dense subspace. Hence $M_{\bar{h}} g = g'$ (in L^2), and $g \in \mathcal{D}(M_{\bar{h}}) = \mathcal{D}(M_h)$. $\square$

Definition 7.6. Let T be a densely defined operator in $\mathcal{H}$. Let

$$\mathcal{D}(T^*) = \{\, h \in H : v \mapsto (Tv \mid h) \text{ is continuous} \,\}.$$

If $h \in \mathcal{D}(T^*)$ then we define $T^* h$ to be the unique vector in $\mathcal{H}$ such that

$$(Tv \mid h) = (v \mid T^* h) \quad \text{for all } v \in \mathcal{D}(T).$$

Example 7.7. As we saw in Example 7.5 on the preceding page, $\mathcal{D}(M_h) = \mathcal{D}(M_h^*)$ and $M_h^* = M_{\bar{h}}$. So if h is real-valued, then $M_h^* = M_h$ and it makes sense to call M_h *self-adjoint*.

Definition 7.8. A densely defined operator T in $\mathcal{H}$ is called *self-adjoint* if $T^* = T$. That is, $\mathcal{D}(T^*) = \mathcal{D}(T)$ and $(Th \mid k) = (h \mid Tk)$ for all $h, k \in \mathcal{D}(T)$.

Example 7.9. Let $\mathcal{H} = L^2(\mathbf{R})$. We want to define $Tf := if'$. It is not unnatural to guess that we should take

$$\mathcal{D}(T) := \{\, f \in L^2(\mathbf{R}) : f' \text{ exists almost everywhere and } f' \in L^2(\mathbf{R}) \,\}.$$

Suppose that $g \in \mathcal{D}(T^*)$ and let $g_0 := T^* g$. If $f \in \mathcal{D}(T)$, then there are *step functions* $\{ f_n \}$ such that $f_n \to f$ in $L^2(\mathbf{R})$.[28] But then

$$\begin{aligned} (Tf \mid g) &= (f \mid g_0) \\ &= \lim_n (f_n \mid g_0) \\ &= \lim_n (Tf_n \mid g) \\ &= 0. \end{aligned}$$

[28] A step function is a linear combination of characteristic functions of bounded intervals.

Since $C_c(\mathbf{R}) \subset T\big(\mathcal{D}(T)\big)$, we must have $g = 0$. Therefore $\mathcal{D}(T^*) = \{0\}$! In particular, $(T, \mathcal{D}(T))$ is not (unitarily equivalent to) a multiplication operator.[29]

Remark 7.10. Suppose that T is any densely defined operator in $\mathcal{H}$. Let $\{h_n\}$ be a sequence in $\mathcal{D}(T^*)$ such that $h_n \to h$ and $T^*h_n \to w$. Let $v \in \mathcal{D}(T)$. Then

$$\begin{aligned}(Tv \mid h) &= \lim_n (Tv \mid h_n) \\ &= \lim_n (v \mid T^*h_n) \\ &= (v \mid w).\end{aligned}$$

Therefore $h \in \mathcal{D}(T^*)$ and $T^*h = w$. In sum, the *graph of* T^*,

$$\mathcal{G}(T^*) := \{ (h, T^*h) \in \mathcal{H} \times \mathcal{H} : h \in \mathcal{D}(T^*) \},$$

is *closed* in $\mathcal{H} \times \mathcal{H}$.

Definition 7.11. We say that an operator T in $\mathcal{H}$ is *closed* if its graph

$$\mathcal{G}(T) := \{ (h, Th) \in \mathcal{H} \times \mathcal{H} : h \in \mathcal{D}(T) \}$$

is closed in $\mathcal{H} \times \mathcal{H}$.

From Remark 7.10, we have the following result.

Lemma 7.12. *If T is any densely defined operator in $\mathcal{H}$, then $(T^*, \mathcal{D}(T^*))$ is a closed operator.*

Definition 7.13. If S and T are operators in $\mathcal{H}$, then we say that T *extends* S if $\mathcal{D}(S) \subset \mathcal{D}(T)$ and $T|_{\mathcal{D}(S)} = S$.

Remark 7.14. Notice that T extends S if and only if $\mathcal{G}(S) \subset \mathcal{G}(T)$. In particular, *if* S has a closed extension, then it has a smallest such.

Definition 7.15. An operator T in $\mathcal{H}$ is called *closable* if it has a closed extension. The smallest such extension is denoted by $\overline{T}$.

The proof of the next result is left as an exercise.

Lemma 7.16. *An operator T in $\mathcal{H}$ is closable if and only if $\overline{\mathcal{G}(T)}$ is the graph of an operator. If T is closable, then $\overline{\mathcal{G}(T)} = \mathcal{G}(\overline{T})$.*

Example 7.17. Let $\mathcal{H} = L^2(\mathbf{R})$ and $T = i\frac{d}{dx}$ be as in Example 7.9 on the facing page. If $f \in \mathcal{D}(T)$, then there are step functions $f_n \to f$ in $L^2(\mathbf{R})$. Therefore $(f, 0) \in \overline{\mathcal{G}(T)}$. But so is (f, f'). Since $\overline{\mathcal{G}(T)}$ is a subspace, $(0, f') \in \overline{\mathcal{G}(T)}$. Therefore $\overline{\mathcal{G}(T)} = \mathcal{H} \times \mathcal{H}$. Thus, T is "way not closable".

Definition 7.18. An operator T in $\mathcal{H}$ is called *symmetric* if $T \subset T^*$.

Example 7.19. A symmetric operator is always closable.

Proposition 7.20. *A densely defined operator T in $\mathcal{H}$ is closable if and only if $\mathcal{D}(T^*)$ is dense. If T is closable, then $\overline{T} = T^{**}$ and $\overline{T}^* = T^*$.*

---

[29]It is interesting to note that if we alter $\mathcal{D}(T)$ so that f is absolutely continuous, then T is self-adjoint [8, §13, Problem 8a]. (Absolute continuity for a function on all of $\mathbf{R}$ is defined in [3].

Proof. Notice that $(h,w)\in\mathcal{G}(T^*)$ if and only if $(Tv\mid h)=(v\mid w)$ for all $v\in\mathcal{D}(T)$. This happens if and only if $\big((-Tv,v)\mid(h,w)\big)=0$ for all $v\in\mathcal{D}(T)$. So we define a unitary operator $V:\mathcal{H}\times\mathcal{H}\to\mathcal{H}\times\mathcal{H}$ by $V(h,k):=(-k,h)$. Then

$$(h,w)\in\mathcal{G}(T^*)\iff(h,w)\in V\big(\mathcal{G}(T)\big)^{\perp}\quad\text{and therefore}\tag{7.1}$$

$$\mathcal{G}(T^*)=V\big(\mathcal{G}(T)\big)^{\perp}.\tag{7.2}$$

Now if $\mathcal{D}(T^*)$ is dense, then T^{**} is defined and

$$\mathcal{G}(T^{**})=V\big(\mathcal{G}(T^*)\big)^{\perp}=V\big(V\big(\mathcal{G}(T)\big)^{\perp}\big)^{\perp}.$$

Since V is unitary and $V^2=-I$, and since we certainly have $\overline{\mathcal{G}(T)}=\mathcal{G}(T)^{\perp\perp}$, you can check that

$$V\big(V\big(\mathcal{G}(T)\big)^{\perp}\big)^{\perp}=\mathcal{G}(T)^{\perp\perp}=\overline{\mathcal{G}(T)}.$$

Therefore $\overline{\mathcal{G}(T)}$ is the graph of an operator, and $\overline{T}=T^{**}$ (Lemma 7.16 on the previous page). We've proved half of the first statement and the first part of the second.

But, assuming $\overline{T}$ exists,

$$\begin{aligned}\mathcal{G}(T^*)&=V\big(\mathcal{G}(T)\big)^{\perp}=V\big(\overline{\mathcal{G}(T)}\big)^{\perp}\\&=V\big(\mathcal{G}(\overline{T})\big)^{\perp}\\&=\mathcal{G}(\overline{T}^{*}).\end{aligned}$$

This proves the second part of the second statement.

Now suppose that $\overline{T}$ exists. Suppose that $w\in\mathcal{D}(T^*)^{\perp}$ and $v\in D(T^*)$. Then

$$\begin{aligned}\big((0,w)\mid V(v,T^*v)\big)&=\big((0,w)\mid(-T^*v,v)\big)\\&=(w\mid v)\\&=0.\end{aligned}$$

Thus $(0,w)\in V\big(\mathcal{G}(T^*)\big)^{\perp}=V\big(V\big(\mathcal{G}(T)\big)^{\perp}\big)^{\perp}=\overline{\mathcal{G}(T)}=\mathcal{G}(\overline{T})$. Therefore, $w=0$ and $\mathcal{D}(T^*)$ is dense. □

As a corollary of (7.2) we have the following.

Corollary 7.21. *Suppose that T and S are operators in $\mathcal{H}$ such that $T\subset S$. Then $S^*\subset T^*$.*

Proof. If $T\subset S$, then $\mathcal{G}(T)\subset\mathcal{G}(S)$. Therefore $V\big(\mathcal{G}(T)\big)\subset V\big(\mathcal{G}(S)\big)$, and we must have $V\big(\mathcal{G}(T)\big)^{\perp}\supset V\big(\mathcal{G}(S)\big)^{\perp}$. Therefore $\mathcal{G}(S^*)\subset\mathcal{G}(T^*)$. □

Proposition 7.22. *Suppose that T is a self-adjoint operator in $\mathcal{H}$. If S is a symmetric operator in $\mathcal{H}$ and if $T\subset S$, then $T=S$.*

Proof. $T\subset S\subset S^*\subset T^*=T$. □

Example 7.23 (Variations on $i\frac{d}{dx}$ — taken from [8, Example 13.4]). Let $\mathcal{H} = L^2([0,1])$. Note that $L^2([0,1]) \subset L^1([0,1])$. Also recall that f is absolutely continuous[30] on $[0,1]$ if f is continuous, f' exists almost everywhere and

$$f(x) = f(0) + \int_0^x f'(t)\,dt \quad \text{for all } x \in [0,1].$$

Furthermore, we have the following.

Let T_k be $i\frac{d}{dx}$ with the domain

$$\mathcal{D}(T_1) = \{\, f \in L^2([0,1]) : f \text{ is absolutely continuous and } f' \in L^2([0,1]) \,\}.$$
$$\mathcal{D}(T_2) = \{\, f \in \mathcal{D}(T_1) : f(0) = f(1) \,\}$$
$$\mathcal{D}(T_3) = \{\, f \in \mathcal{D}(T_2) : f(0) = f(1) = 0 \,\}.$$

Thus $T_3 \subset T_2 \subset T_1$. I claim that

$$T_1^* = T_3, \quad T_2^* = T_2, \quad \text{and} \quad T_3^* = T_1.$$

In particular, each T_k is closed, T_2 is self-adjoint, T_3 is symmetric, and T_1 is not symmetric and has no symmetric extension.

Proof. For the proof, we need this chestnut.

Claim 1. If f and k are absolutely continuous on $[0,1]$, then

$$\int_0^1 f'(x)\overline{k(x)}\,dx + \int_0^1 f(x)\overline{k'(x)}\,dx = f(1)\overline{k(1)} - f(0)\overline{k(0)}.$$

Proof. See [6, §26]. Also, note that $\frac{d}{dx}\big(f(x)k(x)\big) = f'(x)k(x) + f(x)k'(x)$ for almost all x. $\square$

Using this, we see that for f and g absolutely continuous and $f \in \mathcal{D}(T_k)$ we have

$$\begin{aligned}(T_k f \mid g) &= \int_0^1 if'\overline{g} \\ &= -\int_0^1 if\overline{g'} + f(1)\overline{g(1)} - f(0)\overline{g(0)} \\ &= \int_0^1 f\overline{ig'} + f(1)\overline{g(1)} - f(0)\overline{g(0)}.\end{aligned}$$

From this, we quickly deduce that

$$T_3^* \subset T_1, \quad T_2^* \subset T_2, \quad \text{and} \quad T_1 \subset T_3^*. \tag{7.3}$$

Now let $g \in \mathcal{D}(T_k)$, $h = T_k g$ and $H = \int_0^x h$. Then if $f \in \mathcal{D}(T^k)$,

$$\int_0^1 if'g = (T_k f \mid g) = (f \mid h) = f(1)\overline{H(1)} - \int_0^1 f'h.$$

If $k = 1$ or $k = 2$, then $\mathcal{D}(T_k)$ contains the constant functions and we deduce that in both those cases $H(1) = 0$. When $k = 3$, then $f(1) = 0$. Thus in all cases, $(T_k \mid ig) = (T_k f \mid H)$, and

$$ig - H \in \operatorname{Range}(T_k)^\perp. \tag{7.4}$$

[30] I am using a bit of classical measure theory here. For the precise definition of absolute continuity see [7, Chap. 5, §4]. In fact, it would wise to have a look at all of [7, Chap. 5, §3 & §4]. The key results are Theorem 14 and Corollary 15 of §4 of [7, Chap. 5].

If $k = 1$, then Range(T_1) is all of L^2 so we get $ig = H$. Then $ig' = h$ and since $H(1) = 0$, $g \in \mathcal{D}(T_3)$. This shows that $T_1^* \subset T_3$. Combined with (7.3), this shows $T_3 = T_1^*$.

If $k = 2$ or $k = 3$, then Range$(T_k) = \{\, u \in L^2 : \int_0^1 u = 0 \,\}$. Then (7.4) tells us that $ig - H$ is a constant. In particular, we again have $ig' = h$. Furthermore, this also says that g is absolutely continuous. Hence $g \in \mathcal{D}(T_1)$. Thus $T_3^* \subset T_1$. Again, (7.3) now implies equality.

If $k = 0$, we saw earlier that we also have $H(1) = 0$. Hence $g(0) = g(1)$, and $g \in \mathcal{D}(T_2)$. Then $T_2^* \subset T_2$. And we're done. □

Remark 7.24 (Uniqueness). We just proved that the symmetric operator T_3 has a self-adjoint extension — namely T_2. Are there others? The answer appears to be yes, lots. Jody Trout pointed me to [2, Example 2.2.1] as well as some notes online.

7.3. **The Spectrum.** We want to define the *spectrum* $\sigma(T)$ of an operator in $\mathcal{H}$. Again, we look to the case of multiplication operators for guidance.

Example 7.25. Let $h : X \to \mathbf{C}$ be a measurable function and consider the multiplication operator M_h on $L^2(X, \mu)$ (for some Radon measure μ on the locally compact Hausdorff space X). We define the *essential range* of h to be the set of λ such that $\mu\bigl(h^{-1}\bigl(B_\epsilon(\lambda)\bigr)\bigr) > 0$ for all $\epsilon > 0$. Then the essential range is a closed subset of $\mathbf{C}$, and if λ is not in the essential range, then $\frac{1}{\lambda - h}$ is in L^∞ and defines a bounded operator $M_{(\lambda-h)^{-1}}$ on $L^2(X)$.

Definition 7.26. If T is an operator in $\mathcal{H}$, then $\lambda \in \mathbf{C}$ belongs to the *resolvent* $\rho(T)$ *of* T if $\lambda I - T$ is a bijection of $\mathcal{D}(T)$ onto $\mathcal{H}$ and $(\lambda I - T)^{-1} \in B(\mathcal{H})$. The *spectrum* $\sigma(T)$ *of* T is the complement of $\rho(T)$.

Remark 7.27. If $\lambda \in \rho(T)$, then $(\lambda I - T)^{-1}$ is a bounded operator and therefore has closed graph. Then $\lambda I - T$ must also have a closed graph, so T must be a closed operator. Therefore if T is not closed, then $\rho(T) = \emptyset$. On the other hand, if T is a closed operator and if $\lambda I - T$ is a bijection of $\mathcal{D}(T)$ onto $\mathcal{H}$, then $(\lambda I - T)^{-1}$ is necessarily a bounded operator by the closed graph theorem, and $\lambda \in \rho(T)$. (Therefore the condition that $(\lambda I - T)^{-1} \in B(\mathcal{H})$ can be dropped from Definition 7.26 when T is closed to begin with — which is the only case of interest.)

Example 7.28. Suppose that T and S are closed operators in $\mathcal{H}$. In order to define a densely defined operator from their sum, $T + S$, it is necessary to *assume* that $\mathcal{D}(T) \cap \mathcal{D}(S)$ is dense. However, even so, it does not follow that $S + T$ need be closed. For example, let $S = -T$, then $S + T = 0|_{\mathcal{D}(T)}$ which is certainly not closed.[31]

Definition 7.29. If T is a closed operator in $\mathcal{H}$ and if $\lambda \in \rho(T)$, then we write

$$R(\lambda, T) = (\lambda I - T)^{-1}.$$

Theorem 7.30. *If T is a closed operator in $\mathcal{H}$, then $\rho(T)$ is open in $\mathbf{C}$ and $\lambda \mapsto R(\lambda, T)$ is a (strongly) analytic function on $\rho(T)$. Furthermore,*

$$R(\lambda, T) - R(\lambda_0, T) = -(\lambda - \lambda_0) R(\lambda, T) R(\lambda_0, T). \tag{7.5}$$

In particular, $\{\, R(\lambda, T) \,\}_{\lambda \in \rho(T)}$ is a commutative set.

[31] It is worth remarking that a densely defined *bounded* operator can never be closed unless its domain is all of $\mathcal{H}$.

Proof. Note that we trivially have

$$\begin{aligned} R(\lambda, T) &= R(\lambda, T)(\lambda_0 I - T)R(\lambda_0, T), \quad \textit{and} \\ R(\lambda_0, T) &= R(\lambda, T)(\lambda I - T)R(\lambda_0, T). \end{aligned}$$

Subtracting, we get (7.5).

Suppose that $\lambda_0 \in \rho(T)$. For all $|\lambda - \lambda_0| < \|R(\lambda_0, T)\|^{-1}$, we can define

$$S(\lambda, T) := R(\lambda_0, T) \sum_{n=0}^{\infty} (-1)^n (\lambda - \lambda_0)^n R(\lambda_0, T)^n.$$

It is helpful to note that $S(\lambda, T) = (I + (\lambda - \lambda_0)R(\lambda_0, T))^{-1}$.

Clearly, $\lambda \mapsto S(\lambda, T)$ is strongly analytic and $S(\lambda, T)$ maps $\mathcal{H}$ into $\mathcal{D}(T)$. Now using (7.5), you can check that

$$(\lambda I - T)S(\lambda, T) = I \quad \text{and} \quad S(\lambda, T)(\lambda I - T) = I|_{\mathcal{D}(T)}.$$

Therefore, $\lambda \in \rho(T)$ and $S(\lambda, T) = R(\lambda, T)$. This suffices. □

Example 7.31. Let $T = i\frac{d}{dx}$ and let

$$\mathcal{D}(T) = \{\, f \in L^2([0,1]) : f \text{ is absolutely continuous and } f' \in L^2([0,1]) \,\}.$$

Then as we showed in Example 7.23 on page 32, T is a closed operator. However, for any $\lambda \in \mathbf{C}$, $(\lambda I - T)e^{-i\lambda x} = 0$, Therefore $(\lambda I - T)$ is never bijective and $\rho(T) = \emptyset$. (And $\sigma(T) = \mathbf{C}$.)

Example 7.32. With the same set up as Example 7.31, but let

$$\mathcal{D}(T) = \{\, f \in L^2([0,1]) : f \text{ is absolutely continuous, } f' \in L^2 \text{ and } f(0) = 0 \,\}.$$

For each $\lambda \in \mathbf{C}$, define

$$S_\lambda f(x) := \int_0^x e^{-i\lambda(x-t)} f(t)\, dt.$$

Then it is not hard to check that

$$(\lambda I - T)S_\lambda = I \quad \text{and} \quad S_\lambda(\lambda I - T) = I|_{\mathcal{D}(\mathbf{T})}.$$

Therefore in this case $\rho(T) = \mathbf{C}$ and $\sigma(T) = \emptyset$.

Remark 7.33. I'm told that any closed subset of $\mathbf{C}$ can be the spectrum of a closed operator in $\mathcal{H}$ (provided $\mathcal{H}$ is infinite dimensional).

Theorem 7.34. *Suppose that T is a symmetric operator in $\mathcal{H}$. Then the following statements are equivalent.*

(a) *T is self-adjoint.*
(b) *T is closed and for all $\lambda \in \mathbf{C} \setminus \mathbf{R}$,* $\ker(T^* - \lambda I) = \{\, 0 \,\}$.
(c) *T is closed and* $\ker(T^* \pm iI) = \{\, 0 \,\}$.
(d) *For all $\lambda \in \mathbf{C} \setminus \mathbf{R}$,* $\operatorname{Range}(T - \lambda I) = \mathcal{H}$.
(e) $\operatorname{Range}(T \pm iI) = \mathcal{H}$.
(f) $\sigma(T) \subset \mathbf{R}$.

Proof. We'll show that

$$\begin{tikzcd}
 & \text{(b)} \arrow[d, Rightarrow] \arrow[dr, Rightarrow] & \text{(f)} \arrow[l, Rightarrow] & \text{(b) \& (d)} \arrow[l, Rightarrow] \\
\text{(a)} \arrow[ur, Rightarrow] & \text{(c)} \arrow[d, Rightarrow] & \text{(d)} \arrow[dl, Rightarrow] & \\
 & \text{(e)} \arrow[ul, Rightarrow] & &
\end{tikzcd}$$

(a) $\Longrightarrow$ (b): We have $T = T^*$, and T is closed. If $v \in \ker(T - \lambda I)$, then $Tv = \lambda v$ and

$$\lambda(v \mid v) = (Tv|v) = (v|Tv) = \bar{\lambda}(v \mid v).$$

Therefore $v = 0$ if $\lambda \neq \bar{\lambda}$.

(b) $\Longrightarrow$ (c): Easy.

(b) $\Longrightarrow$ (d): Suppose that $v \in \operatorname{Range}((T - \lambda I))^{\perp}$. Then $\big((T - \lambda I)w \mid v\big) = 0$ for all $w \in \mathcal{D}(T)$. Thus $v \in \mathcal{D}((T - \lambda I)^*) - \mathcal{D}(T^*)$. Thus for all $w \in \mathcal{D}(T)$,

$$0 = \big((T - \lambda I)w \mid v\big) = \big(w \mid (T^* - \bar{\lambda})v\big).$$

Since $\mathcal{D}(T)$ is dense, $v \in \ker(T^* - \bar{\lambda} I)$. Since the later is trivial by assumption, $T - \lambda I$ has dense range for all $\lambda \notin \mathbf{R}$.

If $\lambda = a + ib$ and if $v \in \mathcal{D}(T)$, then

$$\begin{aligned}
\big((T - \lambda I)v \mid (T - \lambda I)v\big) &= \big((T - aI)v - ibv \mid (T - aI)v - ibv\big) \\
&= \|(T - aI)v\|^2 + \big((T - aI)v \mid -ibv\big) + \big(-ibv \mid (T - aI)v\big) + |b|^2\|v\|^2.
\end{aligned}$$

Since $(T - aI) \subset (T - aI)^*$, the middle terms cancel, and we have

$$\|(T - \lambda I)v\|^2 = \|(T - aI)v\|^2 + |b|^2\|v\|^2. \tag{7.6}$$

If $w \in \mathcal{H}$, then since the range of $T - \lambda I$ is dense, there must be $v_n \in \mathcal{D}(T)$ such that $(T - \lambda I)v_n \to w$. Then (7.6) (applied to $v_n - v_m$) and $b \neq 0$ implies that $\{ v_n \}$ is Cauchy — say, $v_n \to v$. Since $T - \lambda I$ is a closed operator, $v \in \mathcal{D}(T)$ and $(T - \lambda I)v = w$. This proves (d).

(d) $\Longrightarrow$ (e): Clear.

(e) $\Longrightarrow$ (a): We have $\mathcal{D}(T) \subset \mathcal{D}(T^*)$ by assumption, so we only need to prove the reverse containment. Let $v \in \mathcal{D}(T^*)$. By assumption, there is a $w \in \mathcal{D}(T)$ such that

$$(T - iI)w = (T^* - iI)v.$$

But we also have $w \in \mathcal{D}(T^*)$ so $(T^* - iI)(w - v) = 0$. If $h \in \mathcal{H}$, then there is a $h' \in \mathcal{D}(T)$ with $(T + iI)h' = h$. Therefore

$$\begin{aligned}
(h \mid w - v) &= \big((T + iI)h' \mid w - v\big) \\
&= \big(h' \mid (T^* - iI)(w - v)\big) \\
&= 0.
\end{aligned}$$

Since h is arbitrary, $w = v$, and $w \in \mathcal{D}(T)$. This proves (a).

(f) $\Longrightarrow$ (b): If $\lambda \notin \mathbf{R}$, then $\lambda \in \rho(T)$. Therefore T is closed and $(T - \lambda I)$ is a bijection. In particular, $(T - \lambda I)$ has trivial kernel.

(c) $\Longrightarrow$ (e): Just as in (b) $\Longrightarrow$ (d).

Finally, (b) and (d) together with the closed graph theorem, imply that $\mathbf{C} \setminus \mathbf{R} \subset \rho(T)$. Therefore $\sigma(T) \subset \mathbf{R}$. This completes the proof. □

Theorem 7.35 (Spectral Theorem for Unbounded Self-Adjoint Operators)**.** *Suppose that T is a self-adjoint operator in $\mathcal{H}$. Then there is a locally compact space Y, a Radon measure μ, a continuous function $h : Y \to \mathbf{R}$ and a unitary $U : \mathcal{H} \to L^2(Y, \mu)$ such that*

(a) $U\mathcal{D}(T) = \mathcal{D}(M_h) := \{ f \in L^2(Y, \mu) : hf \in L^2(Y, \mu) \}$ *and*
(b) $UTv = M_h Uv$ *for all* $v \in \mathcal{D}(T)$.

Proof. Theorem 7.34 on page 35 implies that $R(\pm i, T)$ are well-defined bounded operators on $\mathcal{H}$. Since $R(\pm i, T) = (\pm iI - T)^{-1}$, we have

$$\big(R(i, T)(iI - T)v \mid (-iI - T)w\big) = \big((iI - T)v \mid R(-i, T)(-iI - T)w\big).$$

Since Theorem 7.34 also gives us $(\pm iI - T)\mathcal{D}(T) = \mathcal{H}$, we have shows that $R(i, T)^* = R(-i, T)$. Since $R(i, T)$ and $R(-i, T)$ commute (Theorem 7.30 on page 34), $R(i, T)$ is normal. Hence $R(i, T)$ generates a commutative C^*-algebra $A \subset B(\mathcal{H})$. Since $R(i, T)\mathcal{H} = \mathcal{D}(T)$, it follows that the identity representation of A on $B(\mathcal{H})$ is nondegenerate. Thus we can decompose $\mathcal{H}$ into cyclic subspaces as in Section 7.1 on page 28. Therefore we obtain a locally compact space Y and a Radon measure μ such that there is a unitary $U : \mathcal{H} \to L^2(Y, \mu)$ intertwining $R(i, T)$ and M_g for a *bounded* continuous function $g : Y \to \mathbf{C}$.

Since $\ker\big(R(i, T)\big) = \{ 0 \}$, the closed set

$$C := \{ y \in Y : g(y) = 0 \}$$

is a μ-null set. Thus we can replace Y by $Y \setminus C$ (which is still locally compact), and assume that g never vanishes. Let

$$h := i - \frac{1}{g}.$$

Then h is continuous on Y, but may no longer be bounded.

If $v \in \mathcal{D}(T)$, then since $R(i, T)\mathcal{H} = \mathcal{D}(T)$, there is a $w \in \mathcal{H}$ such that $R(i, T)w = v$. Now

$$hUv = hUR(i, T)w = hgUw = (i - \frac{1}{g})gUw = (ig - 1)Uw,$$

and the later is in $L^2(Y, \mu)$. This shows that

$$U\mathcal{D}(T) \subset \mathcal{D}(M_h).$$

Furthermore,

$$\begin{aligned} U(iI - T)U^*Uv &= U(iI - T)U^*UR(i, T)w = Uw \quad \text{and} \\ M_{(i-h)}Uv &= M_{(i-h)}UR(i, T)w = Uw. \end{aligned}$$

Therefore $U(iI - T)U^* \subset M_{(i-h)}$, and $UTU^* \subset M_h$. But then

$$M_h^* = M_{\bar{h}} \subset UT^*U^* = UTU^* \subset M_h.$$

Since $\mathcal{D}(M_h) = \mathcal{D}(M_{\bar{h}})$, we have $UTU^* = M_h = M_{\bar{h}}$. In particular, h must be real-valued and $U\mathcal{D}(T) = \mathcal{D}(M_h)$. □

Corollary 7.36 (Functional Calculus for Unbounded Self-Adjoint Operators)**.** *Suppose that T is a self-adjoint operator in $\mathcal{H}$. Then there is a $*$-homomorphism Φ from the the bounded Borel functions on $\mathbf{R}$ (viewed a C^*-algebra with respect to the supremum norm) into $B(\mathcal{H})$ such that $\Phi(r_\pm) = R(\pm i, T)$, where*

$$r_\pm(x) := \frac{1}{\pm i - x}.$$

Proof. In view of Theorem 7.35 on the previous page, we may as well assume that $T = M_h$ on $L^2(Y, \mu)$ for a continuous function $h : Y \to \mathbf{R}$. Examining the proof of Theorem 7.35 shows that $R(i, T)$ is given by M_g where $g = 1/(i-h)$. Suppose that F is a bounded Borel function on $\mathbf{R}$. Then $F \circ h$ is a bounded Borel function on Y and $M_{F\circ h}$ is a bounded operator on $L^2(Y, \mu)$ with norm bounded by the essential supremum of $F \circ h$. $\Phi(F) = M_{F\circ h}$, then Φ is clearly a $*$-homomorphism.

If we let $F = r_+$, then we have $F \circ h = g$, so $\Phi(r_+) = R(i, T)$. Since $R(-i, T) = R(i, T)^*$, we must also have $\Phi(r_-) = \Phi(\overline{r_+}) = \Phi(r_+)^* = R(-i, T)$, and we are done. □

To conclude this section, we look at the converse of Theorem 7.1 on page 29. This will also serve as another example of the functional calculus at work.

Theorem 7.37 (Stone's Theorem (the rest of the story))**.** *Suppose that T is a self-adjoint operator in $\mathcal{H}$. Then $u_r := e^{irT}$ (that is, $u_r = \Phi(t \mapsto e^{irt})$) defines a one-parameter group of unitaries on $\mathcal{H}$. Moreover*

(a) *$\mathcal{D}(T) = \{\, v \in \mathcal{H} : \lim_{r\to 0} \frac{1}{r}(u_r v - v)$ exists $\}$, and*
(b) *for all $v \in \mathcal{D}(T)$,*

$$\lim_{r\to 0} \frac{u_r v - v}{r} = iTv.$$

Proof. We can assume that $T = M_h$ on $L^2(Y, \mu)$, and hence that $u_r = M_{e^{irh(\cdot)}}$. We certainly see that each u_r is unitary and that $u_{r+s} = u_r u_s$. We need to see that $r \mapsto u_r$ is strongly continuous. But if $f \in L^2(Y)$ and if $r_n \to r$, then

$$e^{ir_n h(y)} f(y) \to e^{irh(y)} f(y)$$

pointwise, and each term is dominated by $|f| \in L^2(Y)$. Hence we get convergence in L^2 via the dominated convergence theorem. (This is a good exercise.)

Now suppose that $f \in \mathcal{D}(T)$. Then

$$\frac{u_r f - f}{r} = \Bigl(\frac{e^{irh(\cdot)} - 1}{r}\Bigr) f(\cdot).$$

There is a $M > 0$ such that

$$\Bigl|\frac{e^{is} - 1}{s}\Bigr| \le M \quad \text{for all } s \in \mathbf{R}.$$

Hence,

$$\Bigl|\frac{e^{irt} - 1}{t}\Bigr| \le Mt \quad \text{for all } r \in \mathbf{R}.$$

But then

$$\Bigl|\frac{e^{irh(y)} - 1}{r} f(y)\Bigr| \le M|h(y)||f(y)|.$$

But $|hf| \in L^2(Y)$. So the dominated convergence theorem implies that

$$\frac{u_r f - f}{r} \to ihf = iT(f) \quad \text{in } L^2(Y).\rtimes$$

Conversely, suppose that

$$\lim_r \frac{u_r f - f}{r} = \lim_r \Bigl(\frac{e^{irh(\cdot)} - 1}{r}\Bigr) f(\cdot) \tag{7.7}$$

converges (in $L^2(Y)$) to g. Since the right-hand side of (7.7) converges pointwise to ihf, we must have $g = ihf$ in $L^2(Y)$. Therefore $hf \in L^2(Y)$ and $f \in \mathcal{D}(T) = \mathcal{D}(M_h)$. $\square$

References

[1] William Arveson, *An Invitation to C^*-algebras*, Springer-Verlag, New York, 1976. Graduate Texts in Mathematics, No. 39.

[2] John B. Conway, *A course in functional analysis*, Graduate texts in mathematics, vol. 96, Springer-Verlag, New York, 1985.

[3] Gerald B. Folland, *Real analysis*, Second, John Wiley & Sons Inc., New York, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[4] James R. Munkres, *Topology: Second edition*, Prentice Hall, New Jersey, 2000.

[5] Iain Raeburn and Dana P. Williams, *Morita equivalence and continuous-trace C^*-algebras*, Mathematical Surveys and Monographs, vol. 60, American Mathematical Society, Providence, RI, 1998.

[6] Frigyes Riesz and Béla Sz.-Nagy, *Functional analysis*, Frederick Ungar, New York, 1955. (English Translation).

[7] Halsey L. Royden, *Real analysis*, Third Edition, Macmillan Publishing Company, New York, 1988.

[8] Walter Rudin, *Functional analysis*, McGraw-Hill Book Co., New York, 1973. McGraw-Hill Series in Higher Mathematics.

[9] ———, *Real and complex analysis*, McGraw-Hill, New York, 1987.

[10] Edwin H. Spanier, *Algebraic topology*, McGraw-Hill, 1966.

[11] Dana P. Williams, *Crossed products of C^*-algebras*, Mathematical Surveys and Monographs, vol. 134, American Mathematical Society, Providence, RI, 2007.

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA

Email address: dana.williams@dartmouth.edu