Spectral Measures

David Loeffler

(Entered for the Yeats Mathematical Essay Prize, October 2003)

Abstract

The theory of spectra for Banach algebras is outlined, including the Gelfand-Naĭmark theorem for commutative B^*-algebras. Resolutions of the identity are introduced, with examples; finally, we prove the spectral theorem for bounded normal operators on a Hilbert space, and conclude with some applications.

1 Definitions

The definitions supplied in the notes for the Lent 2003 part II Functional Analysis course will be assumed.

A **complex Banach algebra** is a complex Banach space A which also posesses a multiplication, with an identity element e having $\|e\| = 1$, satisfying the following properties for all $x, y, z \in A$, $\alpha \in \mathbb{C}$:

$$x(yz) = (xy)z \tag{1}$$
$$(x+y)z = xz + yz, \quad x(y+z) = xy + xz \tag{2}$$
$$\alpha(xy) = (\alpha x)y = x(\alpha y) \tag{3}$$
$$\|xy\| \leq \|x\| \ \|y\| \tag{4}$$

Note that (4) implies that the multiplication is left- and right-continuous (that is, xy is a continuous function of x for every y and a continuous function of y for every x.)

There are two classes of Banach spaces which are sufficiently important to have special notations. One is the space $C(X)$ of continuous complex-valued functions on a nonempty compact Hausdorff space X, with pointwise addition and multiplication and the supremum norm. The other is the space $\mathcal{B}(V)$ of bounded linear operators on a nonempty Banach space V, with the usual operator norm and multiplication defined as composition.

The **spectrum** of $x \in A$ is defined as the set $\sigma(x) = \{\lambda \in \mathbb{C} : \lambda e - x \text{ is not invertible}\}$. The complement of $\sigma(x)$ is known as the **resolvent set**, and the inverse $(\lambda e - x)^{-1}$ is the **resolvent**.

Note that where $A = \mathcal{B}(\mathbb{C}^n)$, the spectrum of $T \in A$ is the set of eigenvalues of T; the spectrum may be regarded as the natural generalisation of eigenvalues to arbitrary Banach algebras.

2 Basic properties of spectra

Theorem 1. *For any $x \in A$, $\sigma(x)$ is a nonempty compact subset of $\mathbb{C}$, and the **spectral radius** $\rho(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}$ has $\rho(x) \leq \|x\| < \infty$.*

Proof. For any $y \in A$ with $\|y\| \leq 1$, $\|y^n\| \leq \|y\|^n$, so the series $\sum_{j=0}^{\infty} (-1)^j y^j$ converges in norm to some $z \in A$. By term-by-term multiplication it is clear that $(e+y)z = z(e+y) = e$. Hence $e + y$ is invertible.

This implies that for $|\lambda| > \|x\|$, $\lambda e - x = \lambda\left(e - \frac{x}{\lambda}\right)$ is invertible, so $\rho(x)$ is finite and bounded above by $\|x\|$. Furthermore, given any invertible $\alpha \in A$, then if $\|\beta\| \leq \|\alpha^{-1}\|^{-1}$, $\alpha + \beta = \alpha\left(e + \alpha^{-1}\beta\right)$ is invertible; so the set G of invertible elements is open in A.

The mapping $R : \lambda \mapsto \lambda e - x$ is evidently continuous, so $R^{-1}(G)$ is open in $\mathbb{C}$; but $R^{-1}(G)$ is the complement of $\sigma(x)$, so $\sigma(x)$ is closed. So it is a closed and bounded subset of $\mathbb{C}$, and therefore compact.

The proof of the nonemptiness of the spectrum clearly must involve some specific properties of $\mathbb{C}$, since it does not hold in real Banach algebras, such as $\mathcal{B}(\mathbb{R}^2)$; there exist rotations of the plane with no real eigenvalues. The following proof (and that of the next theorem) use a deep insight due to Gelfand, which was to study vector spaces by applying classical complex analysis to the linear functionals on them; the form presented here is similar to that in [7].

Suppose that Φ is any bounded linear functional on A. Then $f(\lambda) = \Phi\left[(\lambda e - x)^{-1}\right]$ is an analytic function on $\mathbb{C} \setminus \sigma(x)$, since

$$\lim_{\mu\to\lambda} \frac{f(\lambda) - f(\mu)}{\lambda - \mu} = \lim_{\mu\to\lambda} \frac{\Phi\left[(\lambda e - x)^{-1} - (\mu e - x)^{-1}\right]}{\lambda - \mu} = \lim_{\mu\to\lambda} -\Phi\left[(\lambda e - x)^{-1}(\mu e - x)^{-1}\right] = -\Phi\left[(\lambda e - x)^{-2}\right]$$

exists. (We have used the result that the mapping $x \mapsto x^{-1}$ is continuous on G, which follows from $\|x^{-1} - y^{-1}\| = \|x^{-1}(y-x)y^{-1}\| \leq \|x^{-1}\|\,\|y^{-1}\|\,\|x-y\|$.)

On the other hand, as $\lambda \to \infty$,

$$f(\lambda) = \frac{1}{\lambda}\Phi\left[\left(\frac{x}{\lambda} - e\right)^{-1}\right] \to 0,$$

since $\left(\frac{x}{\lambda} - e\right)^{-1} \to (-e)^{-1} = -e$.

So if $\sigma(x)$ is empty, f is an entire function tending to 0 at infinity, and hence is bounded; so it must be everywhere zero, by Liouville's theorem ([1], ch. 4).

However, let λ_0 be an arbitrary complex number not in $\sigma(x)$. Then $(\lambda_0 e - x)^{-1}$ exists and is nonzero, so by the Hahn–Banach theorem (see [2]) we can find a bounded linear functional Φ_0 such that $\Phi_0\left[(\lambda_0 e - x)^{-1}\right] = 1$, that is $f(\lambda_0) = 1$. This is a contradiction, so the spectrum must be nonempty. □

Theorem 2.

$$\lim_{n\to\infty} \|x^n\|^{1/n} = \rho(x)$$

This may be proved using either the power series approach to complex analysis, or the Cauchy integral formula. We shall use the former approach, following [4]; the latter may be found in [8].

Let $f(\lambda)$ be defined as above. Using the series derived above for $(e+y)^{-1}$, we see that for $\lambda > \|x\|$, the resolvent is given by $(\lambda e - x)^{-1} = \sum_{n=0}^{\infty} \frac{x^n}{\lambda^{n+1}}$, with the sum convergent in the norm topology of A; so as Φ is continuous, f has the Laurent expansion

$$f(\lambda) = \sum_{j=0}^{\infty} \frac{\Phi\left(x^n\right)}{\lambda^{n+1}}$$

It has been shown that f is analytic on the complement of $\sigma(x)$. By the existence and uniqueness theorems for Laurent series, it follows that this sum is in fact convergent for all λ with $\lambda > \rho(x)$. Hence its terms must be bounded; so there exists a constant C (depending on Φ) such that $\Phi\left(x^n\right) \leq C\lambda^n \quad \forall n$.

So the set $\{\frac{x^n}{\lambda^{n+1}} : n \in \mathbb{N}\}$ is **weakly bounded** – that is, every continuous linear functional is bounded on this set. It follows from the Banach–Steinhaus theorem that any weakly bounded set is in fact uniformly bounded; see [2], ch. 5. So there is some constant D such that $\|x^n\| \leq D\lambda^n$, and consequently

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leq \lambda.$$

Since λ was chosen arbitrarily subject to $|\lambda| > \rho(x)$, we conclude that

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leq \rho(x).$$

On the other hand, for any $n > 1$,

$$\lambda^n e - x^n = (\lambda e - x)(\lambda^{n-1}e + \lambda^{n-2}x + \cdots + x^{n-1}).$$

It follows that if $\lambda^n e - x^n$ is invertible, then $\lambda e - x$ must be so; hence $[\rho(x)]^n \leq \rho(x^n) \leq \|x^n\|$. It follows that $\lim_{n\to\infty} \|x^n\|^{1/n}$ exists and is equal to $\rho(x)$. □

One important aspect of this theorem is that if A is a subalgebra of a larger algebra B, then the spectrum σ_A of x considered as an element of A may be strictly larger than its spectrum σ_B as as an element of B, since $\lambda e - x$ may be invertible in B but not in A. However, the spectral radius has been expressed in terms of the norm properties of x alone, which do not depend on the algebra within which x lies.

The following theorem considerably strengthens this result.

Theorem 3. *Every boundary point of $\sigma_A(x)$ lies in $\sigma_B(x)$.*

As before, this may be proved by expressing the algebraic property that λ is a boundary point of $\sigma(x)$ in terms of the norm. The following simple argument is from [3].

Lemma 4. *Let λ_0 be a boundary point of $\sigma(x)$. Then $\|\lambda e - x\|^{-1} \to \infty$ as $\lambda \to \lambda_0$ (through values $\lambda \notin \sigma(x)$).*

Proof. We shall show that if $d(\lambda)$ is the distance from λ to the nearest point of the spectrum, then

$$\|(\lambda e - x)^{-1}\| \geq \frac{1}{d(\lambda)}$$

We know that if y is invertible and $\|z\| < \|y\|$, then $y + z$ is invertible. So $\|(\lambda + \mu)e - x\|$ is invertible whenever $|\mu| < \|\lambda e - x\|$; consequently $d(\lambda) \geq \|\lambda e - x\|$.

By axiom (4) in the definition of a Banach algebra, it is clear that $\|u^{-1}\| \geq \frac{1}{\|u\|}$ for any $u \in A$; so

$$\|(\lambda e - x)^{-1}\| \geq \|\lambda e - x\|^{-1} \geq \frac{1}{d(\lambda)}$$

as required. □

It is clear that any λ having this property is necessarily in $\sigma(x)$, since the resolvent is a continuous function outside $\sigma(x)$. This clearly implies the theorem.

Corollary. *Suppose that the spectrum of $x \in A$ is wholly real. Then every point is a boundary point; so the spectrum of x is the same in any algebra B enclosing A.*

3 Commutative algebras

It is clear that any Banach algebra is a ring; and many of the standard constructions of the theory of commutative rings have analogues for commutative Banach algebras.

A **homomorphism** between Banach algebras is defined to be a map $h : A \to B$, where A and B are Banach algebras, which is linear and satisfies $h(xy) = h(x)h(y)$. An **ideal** is a vector subspace J such that $x \in A, y \in J \Rightarrow xy \in J$. It is clear that the kernel of a homomorphism is an ideal; and conversely, if I is an ideal, then the quotient space A/I has a natural Banach algebra structure which makes the quotient map $A \to A/I$ into a homomorphism.

An ideal is **proper** if it is a proper subset of A, and a proper idea is **maximal** if there is no proper ideal K strictly containing J. A straighforward Zorn's lemma argument shows that any proper ideal lies within some maximal ideal (since the union of any chain of proper ideals ordered by inclusion is an ideal, which is proper as it does not contain e).

For any $x \in A$, the set $[x] = \{xy : y \in A\}$ is evidently an ideal, the **principal ideal** generated by x. This is clearly proper if and only if x is not invertible. If I is any ideal and $x \in I$, then we must have $[x] \subset I$. This implies

Lemma 5. *x is invertible if and only if it it does not lie in any maximal ideal.*

As in the case of rings, ideals in the quotient algebra A/I correspond uniquely to ideals in A containing I. It follows that if I is maximal, A/I has no maximal ideals except $\{0\}$, so every nonzero element is invertible – A/I is a field. This is a very strong requirement to place on a Banach algebra:

Lemma 6 (Gelfand–Mazur). *A Banach algebra which is also a field is isometrically isomorphic to $\mathbb{C}$.*

Proof. By theorem 1 every element x of the algebra must have a nonempty spectrum; but if $\lambda \in \sigma(x)$, $\lambda e - x$ is not invertible and hence must be zero. Thus $x = \lambda e$, as required. □

So for every maximal ideal I there is a (unique) homomorphism from A to $\mathbb{C}$ with kernel I. This implies the following useful characterisation of spectra:

Theorem 7. *Let A be a commutative Banach algebra. The spectrum of x is the set $\{h(x) : h \in \Delta\}$, where Δ is the set of homomorphisms $A \to \mathbb{C}$.*

As any homomorphism is linear, it suffices to prove that x is invertible if and only if there is no $h \in \Delta$ such that $h(x) = 0$. However, this is precisely Lemma 5.

As a corollary, we deduce that for any $h \in \Delta$, $|h(x)| \leq \|x\|$, since $h(x) \in \sigma(x)$.

4 Involutions and B^*-algebras

A B^***-algebra** is a Banach algebra A with a map $x \to x^*$ on A (known as an involution), satisfying the following axioms for all $x, y \in A$, $\alpha \in \mathbb{C}$:

$$(x+y)^* = x^* + y^*$$
$$(\alpha x)^* = \bar{\alpha} x^*$$
$$(xy)^* = y^* x^*$$
$$(x^*)^* = x$$
$$\|xx^*\| = \|x\|^2$$

Note that the fifth axiom implies that $\|x\|\,\|x^*\| \geq \|x\|^2$, or $\|x^*\| \geq \|x\|$; together with the fourth, we see that $\|x^*\| = \|x\|$. If $x^* = x$, x is said to be **hermitian**.

It is clear that for any compact Hausdorff space X, $C(X)$ is a B^*-algebra, with the natural involution $f^*(x) = \overline{f(x)}$. Also, $\mathcal{B}(H)$, the algebra of bounded operators on a Hilbert space, is a B^*-algebra, where M^* is the adjoint of M. (Recall that this is defined by $\langle Mx, y\rangle = \langle x, M^*y\rangle\ \forall x, y \in H$; it may be easily shown that M^* exists and is unique.) That the first four axioms are satisfied is immediate; as for the fifth, on one hand $\|M\|^2 = \|M\|\,\|M^*\| \geq \|MM^*\|$, while on the other hand

$$\|M\|^2 = \|M^*\|^2 = \sup_{\|x\|=1} \langle M^*x, M^*x\rangle = \sup_{\|x\|=1} \langle MM^*x, x\rangle \leq \sup_{\|x\|=1} \|MM^*x\|\,\|x\| = \|MM^*\|,$$

from which it follows that $\|M\|^2 \leq \|MM^*\|$; hence $\|M\|^2 = \|MM^*\|$.

A homomorphism between two B^*-algebras which preserves the involution ($h(x^*) = [h(x)]^*$) is said to be a *-homomorphism. The next theorem claims that any homomorphism from a B^*-algebra to $\mathbb{C}$ is in fact a *-homomorphism.

Theorem 8 (Gelfand–Naĭmark). *Let A be a commutative B^*-algebra. Then for any homomorphism $h : A \to \mathbb{C}$ and any $x \in A$, $h(x^*) = \overline{h(x)}$.*

Proof. It is clear from the definition of an involution that every element may be expressed uniquely as a sum $u + iv$ where u and v are hermitian. Hence it suffices to show that if x is hermitian, then $h(x)$ is real.

Define $z = x + ite$ for $t \in \mathbb{R}$. Then $z^* = x - ite$, so $zz^* = x^2 + t^2e^2$. If $h(x) = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$, then $h(z) = \alpha + i(\beta + t)$; but by Theorem 7,

$$|h(z)|^2 = \alpha^2 + (\beta+t)^2 \leq \|z\|^2 = \|zz^*\| \leq \|u\|^2 + t^2$$

So $\alpha^2 + \beta^2 + 2\beta t \leq \|u\|^2$ for all t, whence $\beta = 0$, as required. □

Corollary. *For any $x \in A$, we have $\|x\| = \rho(x)$.*

Proof. Let $y = xx^*$. Then y is hermitian. By the definition of a B^*-algebra, we have $\|y^2\| = \|yy^*\| = \|y\|^2$.

It follows that $\|y^{2^k}\| = \|y\|^{2^k}$ for all k. Hence

$$\rho(y) = \limsup_{n\to\infty} \|y^n\|^{1/n} = \|y\|.$$

However, for any $h \in \Delta$, $h(y) = h(xx^*) = h(x)h(x^*) = h(x)\overline{h(x)} = |h(x)|^2$. Thus

$$\rho(y) = \max\{|h(y)| : h \in \Delta\} = \max\{|h(x)|^2 : h \in \Delta\} = [\rho(x)]^2$$

However, we know that $\|y\| = \|xx^*\| = \|x\|^2$; so we have $\|x\| = \rho(x)$. □

B^*-algebras have a useful regularity property related to Theorem 3.

Lemma 9. *If $N \in A$ is invertible in $\mathcal{B}(H)$, it is invertible in A.*

Proof. Consider the operator $R = NN^*$. This is self-adjoint, and hence normal; so its spectrum is wholly real. Also, since $(N^*)^{-1} = (N^{-1})^*$, it is invertible in $B(H)$.

If R is not invertible in A, that is $0 \in \sigma_A(R)$, then 0 must be a boundary point of $\sigma_A(R)$, since $\sigma_A(R) \subset \mathbb{R}$. So by Theorem 3, 0 is in the spectrum of N with respect to any superalgebra of A, in particular $\mathcal{B}(H)$. But we know N is invertible in $\mathcal{B}(H)$, which is a contradiction; so N is invertible in A.

Since $M^{-1} = M^*(MM^*)^{-1}$, M is thus invertible in A. □

Corollary. *If A is a closed subalgebra of a B^*-algebra B and A is closed under the involution, then $\sigma_A(x) = \sigma_B(x)$ for all $x \in A$.*

5 Resolutions of the identity

As observed above, $\mathcal{B}(H)$, where H is a Hilbert space, is a B^*-algebra, and the theory developed above may be applied to it to obtain some remarkable representation theorems for normal operators (operators which commute with their adjoint). Henceforth we shall reserve the letters $x, y, \ldots$ for elements of H, and denote elements of $\mathcal{B}(H)$ by $M, N, \ldots$.

Let Ω be a compact subset of $\mathbb{C}$, and $\mathfrak{M}$ the σ-algebra of Borel subsets of Ω. Then a **resolution of the identity** is a map $E : \mathfrak{M} \to \mathcal{B}(H)$ such that

(i) $E(\emptyset) = 0$, $E(\Omega) = e$.

(ii) Each $E(\omega)$ is self-adjoint.

(iii) $E(\omega_1 \cap \omega_2) = E(\omega_1)E(\omega_2)$.

(iv) For every $x \in H$ and $y \in H$, the function $E_{x,y}(\omega) = \langle E(\omega)x, y\rangle$ is a regular complex Borel measure on $\mathfrak{M}$.

Observation 1. Each $E(\omega)$ is necessarily a projection, as $E(\omega)^2 = E(\omega \cap \omega) = E(\omega)$, by (ii).

Observation 2. If M and N are operators on H and $\langle Mx, y\rangle = \langle Nx, y\rangle \ \forall x, y \in H$, then clearly we must have $M = N$; so (iv) and the additivity property of measures implies that we must have

$$E(\omega_1 \cup \omega_2) = E(\omega_1) + E(\omega_2)$$

for any disjoint sets ω_1, ω_2.

Observation 3. For any $x \in H$, $E_{x,x}$ is a positive measure of total variation $\|x\|^2$. This is because $E_{x,x}(\omega) = \langle E(\omega)x, x\rangle = \langle E(\omega)^2(x), x\rangle = \|E(\omega)x\|^2$, as $E(\omega)$ is self-adjoint and idempotent.

Observation 4. The requirement that each $E_{x,y}$ be regular is in fact automatically true, due to theorem 2.18 of [7], which states that on a locally compact Hausdorff space for which every open set is σ-compact, such as any compact subset of $\mathbb{C}$, a Borel measure μ with $\mu(K) < \infty$ for every compact K is necessarily regular.

Let us consider some examples of resolutions of the identity.

Let M be a normal linear operator on $\mathbb{C}^n$, so its spectrum is the set of its eigenvalues $\lambda_1, \ldots, \lambda_m$. Since a normal operator is unitarily diagonalizable ([6], §15.12), the eigenspaces $\mathcal{E}_{\lambda_j}$ are mutually orthogonal and their direct sum is the whole space.

For subsets $\omega \subset \{\lambda_1, \ldots, \lambda_m\}$, define $\tilde{\mathcal{E}}_\omega = \bigoplus_{\lambda \in \omega} \mathcal{E}_\lambda$. If we define $E(\omega)$ to be the orthogonal projection onto $\tilde{\mathcal{E}}_\omega$, then it is easily checked that (i)—(iv) are satisfied.

Since $\langle E(\cdot)x, y\rangle$ is a measure, it makes sense to consider integrals with respect to that measure. We see that

$$\int_\Omega \lambda dE_{x,y}(\lambda) = \sum_{j=1}^{k} \langle \lambda_j E(\{\lambda_j\})x, y\rangle = \left\langle \left[\sum_{j=1}^{k} \lambda E(\{\lambda_j\})\right](x), y \right\rangle$$

What is $\sum_{j=1}^k \lambda E(\{\lambda_j\})$? Considering the effect of M on a basis of eigenvectors, it is easily seen that $\sum_{j=1}^k \lambda E(\{\lambda_j\}) = M$, and

$$\int_\Omega \lambda dE_{x,y}(\lambda) = \langle Mx, y\rangle$$

so this integral reconstructs M from E; we have expressed the diagonalisation of M using the formalism of resolutions of the identity.

The next example illustrates that resolutions of the identity may be used to extend the classical concept of diagonalisability to a much wider class of operators.

Let $L^2[0,1]$ be the Hilbert space of square-integrable complex functions on $[0,1]$, with the inner product

$$\langle x, y\rangle = \int_{[0,1]} x(t)\overline{y(t)}d\mu(t)$$

where μ is Lebesgue measure.

Define the operator M by $[Mx](t) = t\,x(t)$. This is clearly a bounded linear operator of norm 1; it is also self-adjoint, and thus normal. It is not difficult to see that if $\lambda \notin [0,1]$, then $\lambda I - M$ is invertible; on the other hand, if $\lambda \in (0,1)$, then the functions f_n which are n on $[\lambda - \frac{1}{2n}, \lambda + \frac{1}{2n}]$ and 0 otherwise satisfy $\|f_n\| = 1\ \forall n$ and $\|(\lambda I - M)f_n\| = \frac{1}{12n} \to 0$, so $\lambda I - M$ has no bounded inverse. Hence $\sigma(M) = [0,1]$, as $\sigma(M)$ must be closed.

We shall construct a resolution of the identity by the rule

$$[E(\omega)x](t) = \begin{cases} x(t) & (t \in \omega) \\ 0 & (t \notin \omega) \end{cases}$$

Then $E_{x,y}(\omega) = \int_\omega x(t)\overline{y(t)}d\mu(t)$, so $E_{x,y}$ is the weighted measure $x\bar{y}\mu$. This is clearly a complex measure, and satisfies

$$\int_{[0,1]} \lambda dE_{x,y}(\lambda) = \int_{[0,1]} \lambda x(\lambda)\overline{y(\lambda)}d\mu(\lambda) = \langle Mx, y\rangle.$$

The important aspect of this example is that M has no eigenvalues: there is no square-integrable function with $tx(t) = \lambda x(t)$ except the zero function. So it cannot possibly have a diagonalisation in the sense of standard linear algebra.

However, for any closed $V \subset \sigma(M)$ such that $E(V) \neq 0$, M maps the range R of $E(V)$ into itself, so the restriction $M|R$ is a bounded linear operator on R; and $\sigma(M|R) = V$ (that is, the spectrum of M regarded as an element of $\mathcal{B}(R)$ is V.) Roughly speaking, for each $\lambda \in \sigma(M)$, there are spaces which come arbitrarily close to being eigenspaces with eigenvalue λ, but no exact eigenspace.

The purpose of resolutions of the identity is to provide a framework sufficiently flexible to describe such phenomena. It is this statement of the fact that "a normal operator is diagonalizable" which generalises in the most natural way: the *spectral theorem for bounded normal operators* states that for any bounded normal operator M on a Hilbert space, there is a unique resolution of the identity E on the Borel subsets of its spectrum such that

$$\int_{\sigma(M)} \lambda dE_{x,y}(\lambda) = \langle Mx, y\rangle$$

6 Proof of the spectral theorem

The following proof of the spectral theorem, which is similar to that of [3], uses the *Riesz representation theorem* ([5], §36.6; [7], §6.19). This states that if X is a locally compact Hausdorff space and Γ is a bounded linear functional on $C(X)$, there exists a unique regular complex Borel measure μ of total variation $\|\Gamma\|$ such that

$$\Gamma(f) = \int_X f d\mu$$

for all $f \in C(X)$.

The key to the proof is to find a mapping Ψ between $C(\sigma(M))$ and a closed subalgebra of $\mathcal{B}(H)$ such that the identity function on $C(\sigma(M))$ maps to the operator M itself; then for any $x, y \in H$, $f \mapsto \langle \Psi(f)x, y\rangle$ is a linear functional on $C(\sigma(M))$, and the Riesz theorem provides a Borel measure $\mu_{x,y}$ representing this functional, which we shall take to be $E_{x,y}$.

The easiest way to construct such a mapping is to use the B^*-algebra structure of $C(\sigma(M))$, and find a map Ψ preserving as much of this structure as possible: an isometric *-isomorphism. If the identity function $\lambda \mapsto \lambda$

is mapped to M, and the conjugation function $\lambda \mapsto \bar{\lambda}$ to the adjoint M^*, then any polynomial $P(\lambda, \bar{\lambda})$ must be mapped to $P(M, M^*)$, and this map is clearly a homomorphism.

However, $\sigma(M)$ is compact; so by the Stone–Weierstrass theorem ([2], ch. 6) the polynomials $P(\lambda, \bar{\lambda})$ are dense in $C(\sigma(M))$. Hence if the linear mapping Ψ defined above on the subalgebra of polynomials can be shown to be continuous, then it must have a unique continuous extension to all of $C(\sigma(M))$, the range of which must be contained in the closure A of the space of polynomials $P(M, M^*)$ – the smallest closed subalgebra of $\mathcal{B}(H)$ containing M and M^*.

In fact, Ψ is not only continuous, but isometric. Since M is normal, A is a commutative B^*-algebra[1], and the corollary to Theorem 8 implies that

$$\|P(M, M^*)\| = \max\{|\lambda| : \lambda \in \sigma(p(M, M^*))\}.$$

What is $\sigma(P(M, M^*))$? The answer is about as nice as we could hope for.

Lemma 10 (Spectral mapping theorem for polynomials).

$$\sigma[p(M, M^*)] = \{p(\lambda, \bar{\lambda}) : \lambda \in \sigma(M)\}$$

Proof.

$$\begin{aligned}
\sigma(P(M, M^*)) &= \{h(P(M, M^*)) : h \in \Delta\} && \text{(by Theorem 7)} \\
&= \{P(h(M), h(M^*)) : h \in \Delta\} && \text{(since } h \text{ is a homomorphism)} \\
&= \{P(h(M), \overline{h(M)}) : h \in \Delta\} && \text{(by Theorem 8)} \\
&= \{P(\lambda, \bar{\lambda}) : \lambda \in \sigma(M)\}
\end{aligned}$$

□

Therefore,

$$\begin{aligned}
\|P(M, M^*)\| &= \max\{|\lambda| : \lambda \in \sigma(P(M, M^*))\} \\
&= \max\{|P(\lambda, \bar{\lambda})| : \lambda \in \sigma(M)\} \\
&= \|P\|_\infty
\end{aligned}$$

That is, the norm of $P(M, M^*)$ as an element of $\mathcal{B}(H)$ is exactly the norm of P as an element of $C(\sigma(M))$. Hence Ψ is an isometric *-isomorphism, as claimed.

(There is a very good reason why this mapping should be so remarkably regular; it is an example of the *Gelfand transform*, which states that any commutative B^*-algebra is isometrically *-isomorphic to $C(\Delta)$, where Δ is the set of its maximal ideals endowed the weak-* topology. This, however, is outside the scope of this essay; a thorough account is given in [8], ch. 11.)

Using this isomorphism Ψ, it is possible to complete the proof of the spectral theorem.

Theorem 11 (Spectral theorem for bounded normal operators). *For any normal $M \in \mathcal{B}(H)$, there exists a unique resolution of the identity E such that*

$$\int_{\sigma(M)} \lambda dE_{x,y}(\lambda) = \langle Mx, y \rangle.$$

This has the following properties:

(a) For any open set $\omega \in \mathfrak{M}$, $E(\omega) \neq 0$.

(b) Every element of A commutes with each of the projections $E(\omega)$.

[1] Routine verification shows that the closure of a commutative algebra is commutative.

Proof. Define $\mu_{x,y}$ to be the unique regular complex Borel measure on σ representing the functional $\langle\Psi(\cdot)x, y\rangle$, provided by the Riesz theorem.

Then for any Borel subset ω of σ, $\mu_{x,y}(\omega)$ is uniquely defined by ω, x and y. So it must be sesquilinear in x and y for each fixed ω. Also, we must have $|\mu_{x,y}(\omega)| < \|x\|\ \|y\|$: $|\mu_{x,y}(\omega)| < \operatorname{var}\mu_{x,y}$, the total variation of $\mu_{x,y}$ is the norm of the functional $\langle\Psi(\cdot)x, y\rangle$, and $|\langle\Psi(f)x, y\rangle| \leq \|\Psi(f)x\|\ \|y\| \leq \|x\|\ \|y\|\ \|f\|_\infty$, so the norm of this functional is at most $\|x\|\ \|y\|$ as claimed.

Thus for any ω, there exists a unique bounded linear operator $E(\omega)$ such that $\mu_{x,y}(\omega) = \langle E(\omega)x, y\rangle$.

Now, I claim that these $E(\omega)$ satsify the axioms for a resolution of the identity. Let's verify the axioms given above.

(i) $E(\emptyset) = 0, E(\Omega) = I$.

The first requirement is trivial. For the second, $\langle E(\Omega)x, x\rangle = \mu_{x,x}(\sigma(M))$. Since $\mu_{x,x}$ is a positive measure, this is precisely the norm of the functional $\langle\Psi(\cdot)x, x\rangle$. This is clearly bounded above by $\|x\|^2$, and since $\Psi(1) = e$, this bound is attained. Hence $\langle E(\sigma)x, x\rangle = \|x\|^2\ \forall x, y \in H$, and it follows that $E(\sigma(M))$ is the identity.

(ii) *E is self-adjoint.*

The definition of E implies that

$$\langle E(\omega)^*x, y\rangle = \langle x, E(\omega)y\rangle = \overline{\langle E(\omega)y, x\rangle}$$

So we seek to prove that

$$\langle E(\omega)x, y\rangle = \overline{\langle E(\omega)y, x\rangle}$$

or equivalently

$$\mu_{x,y}(\omega) = \overline{\mu_{y,x}}(\omega)$$

Now

$$\int_\sigma f d\overline{\mu_{y,x}} = \overline{\int_\sigma \bar{f} d\mu_{y,x}} = \overline{\langle\Psi(\bar{f})y, x\rangle} = \langle x, \Psi(\bar{f})y\rangle$$

The isomorphism Ψ was defined in such a way that $\Psi(\bar{f}) = (\Psi(f))^*$, so we have shown that

$$\int_\sigma f d\overline{\mu_{y,x}} = \int_\sigma f d\mu_{x,y}$$

for all continuous f, from which it follows that $\overline{\mu_{y,x}} = \mu_{x,y}$ as required.

(iii) $E(\omega_1)E(\omega_2) = E(\omega_1 \cap \omega_2)$.

An appealing proof of this is to use characteristic functions: $E(\omega) = \Psi(\chi_\omega)$, and since Ψ is multiplicative, we have

$$E(\omega_1 \cap \omega_2) = \Psi(\chi_{\omega_1\cap\omega_2}) = \Psi(\chi_{\omega_1}\chi_{\omega_2}) = \Psi(\chi_{\omega_1})\Psi(\chi_{\omega_2}) = E(\omega_1)E(\omega_2)$$

Unfortunately, as written this is not valid, since Ψ is defined on the Banach algebra $C(\sigma)$ of continuous functions, and characteristic functions aren't continuous. However, in the equation

$$\langle\Psi(f)x, y\rangle = \int_\sigma f(\lambda)d\mu_{x,y}(\lambda)$$

the right-hand side is still defined when f is any bounded Borel function on σ, and defines a bounded linear operator on H. Hence this may be used to extend the definition of Ψ to such functions; if we can now show that Ψ is still multiplicative on this larger domain, the desired result will follow.

If f and g are continuous, and $S = \Psi(f)$, $T = \Psi(g)$, then

$$\int_\sigma fg d\mu_{x,y} = \langle STx, y\rangle = \int_\sigma f d\mu_{Tx,y}$$

This holds for every $f \in C(\sigma)$, so $\mu_{Tx,y}$ and $g \cdot \mu_{x,y}$ are equal as measures. Therefore this equality remains true for arbitrary bounded Borel functions f. (This is essentially the uniqueness assertion of the Riesz theorem.)

Repeating the process, and defining $z = (\Psi(f))^* y$, we can reason that for continuous g and arbitrary bounded Borel f, we have

$$\int_\sigma fg d\mu_{x,y} = \int_\sigma f d\mu_{Tx,y} = \langle \Psi(f)Tx, y\rangle = \langle Tx, z\rangle = \int_\sigma g d\mu_{x,z}$$

and we see that the measures $f \cdot \mu_{x,y}$ and $\mu_{x,z}$ are the same, so these formulae hold for all bounded Borel g.

We can now conclude that for arbitrary bounded Borel functions f and g, we have

$$\langle \Psi(fg)x, y\rangle = \int_\sigma fg d\mu_{x,y} = \int_\sigma g d\mu_{x,z} = \langle \Psi(g)x, z\rangle = \langle \Psi(g)x, (\Psi(f))^* y\rangle = \langle \Psi(f)\Psi(g)x, y\rangle$$

and the result follows.

(iv) *For every $x \in H$ and $y \in H$, the function*

$$E_{x,y}(\omega) = \langle E(\omega)x, y\rangle$$

is a complex measure on $\mathfrak{M}$.

This is immediate from our definition of E.

It remains is to show that the resolution of the identity thus constructed is unique. This, however, is immediate, since for each x, y the measure $\mu_{x,y}$ is unique. Hence if there existed two distinct resolutions of the identity E, E', for any $\omega \in \mathfrak{M}$ we would have $\langle E(\omega)x, y\rangle = \langle E'(\omega)x, y\rangle$ for all $x, y \in H$, so $E(\omega) = E'(\omega)$.

The fact that $E(U) \neq 0$ for open sets $U \in \mathfrak{M}$ follows from the fact that if U is open, there exists a continuous function f on $\sigma(M)$ having support in U which is not everywhere zero. If $E(U) = 0$, then

$$\langle \Psi(f)x, y\rangle = \int_{\sigma(M)} f(\lambda) dE_{x,y}(\lambda) = 0 \ \forall x, y \in H,$$

but Ψ is an isometry, and hence injective, which is a contradiction.

To show that each E commutes with every element of A, we note that $E(\omega) = \Psi(\chi_\omega)$. It has been shown that $\Psi(f)\Psi(g) = \Psi(fg)$ for any bounded Borel f and g; it clearly follows that $\Psi(f)\Psi(g) = \Psi(g)\Psi(f)$, and it follows that every E commutes with every element of A. □

The next theorem provides an interpretation of the spectral decomposition in terms of eigenspaces.

Theorem 12. *Let $V \in \mathfrak{M}$. Then if the range R of the spectral projection $E(V)$ is nonempty, it is an invariant subspace of M, and $\overline{\operatorname{Int} V} \subset \sigma(M|R) \subset \overline{V}$ (where $\overline{X}$ is the closure of X).*

Proof. Since $E(V)$ commutes with M, for every $x \in R$ we must have $Mx = ME(V)x = E(V)Mx$, so $Mx \in R$. Thus R is an invariant subspace and the restriction $M|R$ is well-defined.

For any $\mu \notin \overline{V}$, the function $f(\lambda) = \dfrac{\chi_V(\lambda)}{\mu - \lambda}$ is a bounded Borel function on $\sigma(M)$, since it is bounded above by $\frac{1}{d(\mu,V)}$; hence $\Psi(f)$ is a bounded linear operator on H. However, if $g(\lambda) = \mu - \lambda$, then $\Psi(g) = \mu I - M$; and $f(\lambda)g(\lambda) = \chi_V(\lambda)$, so $\Psi(f)\Psi(g) = \Psi(fg) = E(V)$. Since $E(V)$ is a projection, its restriction to its own range R is the identity. Thus $\Psi(f)$ is the inverse of $\mu I - M$ in the algebra $\mathcal{B}(R)$; so $\mu \notin \sigma(M|R)$.

Conversely, suppose $\mu \in \operatorname{Int} V$. Then μ is disjoint from the closure of $\sigma(M) \setminus V$, so by the same logic as above, $\mu I - M$ is invertible on the orthogonal complement S of R, which is the range of $E(\sigma(M) \setminus V)$. Hence if $\mu I - M$ is invertible on R, it is invertible on $R \oplus S = H$; but $V \subset \sigma(M)$, so $\mu I - M$ cannot be invertible on H. Hence it is not invertible on R, so $\mu \in \sigma(M|R)$. Since $\sigma(M|R)$ is a closed set containing $\operatorname{Int} V$, it contains $\overline{\operatorname{Int}(V)}$. □

If we apply this to the singleton closed set $\{\lambda_0\}$, we see that if $R = E(\{\lambda_0\})$ is not zero, then $\sigma(M|R) \subset \{\lambda_0\}$. Since $\sigma(M|R) \neq \emptyset$, it follows that $\sigma(M|R) = \{\lambda_0\}$. Applying the spectral theorem to $M|R$, we see that the functions $f(\lambda) = \lambda$ and $g(\lambda) = \lambda_0$ are equal on $\sigma(M|R)$; hence on $\sigma(M|R)$, we have $M = \lambda_0 I$; so R is an eigenspace of M with eigenvalue λ_0.

Furthermore, if λ_0 is an isolated point of the spectrum, then $\{\lambda_0\}$ is an open set in the toplogy of $\sigma(M)$; so $E(\{\lambda_0\})$ is not empty, and it follows that λ_0 is an eigenvalue.

7 Some consequences of the spectral theorem

Evidently the spectral theorem gives us a great deal of insight into the structure of normal operators on Hilbert space. Let's use this to prove some interesting theorems.

Theorem 13. *A normal operator M is*

(a) self-adjoint if and only if $\sigma(M)$ is real

(b) unitary if and only if $\sigma(M)$ is in the unit circle

Proof. Since the map Ψ is an isometry on $C(\sigma(M))$, it is injective thereon. Thus $\Psi(f) = \Psi(g)$ if and only if $f(\lambda) = g(\lambda)$ for all $\lambda \in \sigma(M)$.

(a) Taking $f(\lambda) = \lambda$ and $g(\lambda) = \bar{\lambda}$, it follows that $\Psi(f) = M$ and $\Psi(g) = M^*$ are equal if and only if $\lambda = \bar{\lambda}$ on $\sigma(M)$, i.e. if and only if $\sigma(M) \subset \mathbb{R}$.

(b) Taking $f(\lambda) = \lambda\bar{\lambda}$ and $g(\lambda) = 1$, we see that $\Psi(f) = MM^*$ and $\Psi(g) = e$ are equal if and only if $|\lambda| = 1 \ \forall \lambda \in \sigma(M)$.

□

For the following theorem, we shall define an operator P to be **positive** if $\langle Px, x\rangle \geq 0$ for all $x \in H$. It is easily seen that any positive operator is self-adjoint, and thus normal.

Theorem 14. *A necessary and sufficient condition for P to be positive is that $\sigma(P) \subset [0, \infty)$.*

The necessity result is elementary (see [8], thm. 12.32). The sufficiency follows from the spectral representation:

$$\langle Px, x\rangle = \int_{\sigma(P)} \lambda d\mu_{x,x}(\lambda) \geq 0$$

since $\mu_{x,x}$ is a positive measure and $\lambda > 0$.

For the next few results we shall need to extend the spectral mapping theorem to a wider class of functions: the spectrum of $\Psi(f)$ is $f(\sigma(M))$, where f is any continuous function.

Proof. We shall first show that $\sigma(\Psi(f)) \subset f(\sigma(M))$; that is to say, if there is no $\lambda \in \sigma(M)$ such that $f(\lambda) = \mu$, then $\mu I - \Psi(f)$ is invertible. However, this is immediate, since $g(\lambda) = \frac{1}{\mu - f(\lambda)} \in C(\sigma(M))$; so $\Psi(g)(\mu I - \Psi(f)) = \Psi(g)\Psi(\mu I - f) = \Psi(1) = e$. So $\mu \notin f(\sigma(M)) \Rightarrow \mu \notin \sigma(\Psi(f))$.

For the second part, we must show that $\Psi(f(\lambda_0) - f(\cdot))$ is not invertible for $\lambda_0 \in \sigma(M)$. If it has an inverse in the minimal closed subalgebra A generated by e, M and M^*, then this must be $\Psi(g)$ for some $g \in C(\sigma(x))$; we then must have $g(\lambda)(f(\lambda_0) - f(\lambda)) = 1$ on $\sigma(M)$, which is clearly not possible. However, by Lemma 9, this is sufficient, since A is a closed subalgebra containing $\Psi(f(\lambda_0) - f(\cdot))$ and its adjoint. □

Theorem 15. *A positive operator P has a positive square root.*

This is immediate, as the principal branch of the square root function is defined and continuous on $\sigma(P)$ and maps it into a subset of the positive real line. Hence $S = \Psi(\sqrt{\cdot})$ exists and satisfies $S^2 = P$; and by the spectral theorem its spectrum lies in $[0, \infty)$, so S is positive. (It may in fact be shown that S is unique, but this requires a little more machinery.)

Theorem 16. *Any invertible bounded operator T may be expressed as UP where U is unitary and P is positive. T is normal if and only if $UP = PU$.*

T^*T is a positive operator, since $\langle T^*Tx, x\rangle = \|T^*x\|^2 \geq 0$. Hence it has a square root P. Since T is invertible, T^*T and P are so.

Hence if we define $U = TP^{-1}$, then $U^*U = P^{-1}T^*TP^{-1} = P^{-1}P^2P^{-1} = I$, so U is unitary, and we have the required decomposition.

It is clear that if $UP = PU$, then $T^* = PU^{-1}$ evidently commutes with T. Conversely, suppose T is normal; then we may define bounded Borel functions $u(\lambda)$ and $p(\lambda)$ on $\sigma(T)$ such that $|u| = 1$, p is real and positive and $u(\lambda)p(\lambda) = \lambda$. ($p$ is clearly continuous everywhere, but u may fail to be continuous at 0.) Their associated operators $U = \Psi(u)$ and $P = \Psi(p)$ have the required properties.

This in turn allows us to prove a striking theorem about the topological properties of the group of invertible operators.

Theorem 17. *The subgroup of invertible operators in* $\mathcal{B}(H)$ *is path-connected.*

We shall show that any invertible operator may be connected to the identity by a continuous arc.

Suppose $T = UP$. Then the spectrum of U is a subset of the unit circle; hence there is a real bounded Borel function f such that $\exp(if(\lambda)) = \lambda$ on $\sigma(U)$. Put $Q = \Psi_U(f)$. Then we have $\exp(iQ) = U$, and Q is self-adjoint.

It follows that $r \mapsto U_r = \exp(irQ)$ is a continuous map of $[0, 1]$ into the group of invertible operators with $U_0 = I$ and $U_1 = U$.

Now, we may straightforwardly connect P to the identity, since the path $r \mapsto rP + (1 - r)I$ clearly passes through only positive invertible operators. Hence $r \mapsto U_rP_r$ is a path from the identity to T inside the set of invertible operators, proving the required result.

References

[1] L. V. Ahlfors. *Complex Analysis.* McGraw–Hill, New York, third edition, 1978.

[2] B. Bollobás. *Linear Analysis.* Cambridge University Press, 1990.

[3] H. R. Dowson. *Spectral Theory of Linear Operators.* Academic Press, London, 1978.

[4] N. Dunford and J. T. Schwartz. *Linear Operators.* Wiley Interscience, New York, 1971.

[5] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis.* Dover Publications, New York, 1975.

[6] S. Lang. *Algebra.* Addison–Wesley, third edition, 1993.

[7] W. Rudin. *Real and Complex Analysis.* McGraw–Hill, New York, third edition, 1987.

[8] W. Rudin. *Functional Analysis.* McGraw–Hill, New York, second edition, 1991.