

8

The Spectral Theorem for Bounded Self-Adjoint Operators: Proofs

In this chapter we give proofs of all versions of the spectral theorem stated in the previous chapter.

8.1 Proof of the Spectral Theorem, First Version

A proof of the spectral theorem, in its projection-valued measure form, can be obtained in two main stages. The first stage of the proof is to define a *continuous* functional calculus, meaning we associate with each continuous function f on $\sigma(A)$ an operator $f(A)$. The map $f \mapsto f(A)$ should have the property that if f is the function $f(\lambda) = \lambda^m$, then $f(A) = A^m$. The continuous functional calculus is then constructed by approximating continuous functions on $\sigma(A)$ by polynomials. The Stone–Weierstrass theorem tells us that polynomials are dense in the continuous functions on $\sigma(A)$; it remains only to show that if a sequence p_n of polynomials converges uniformly to some continuous function f on $\sigma(A)$, then the operators $p_n(A)$ converge to some operator, which we will then call $f(A)$.

The second stage of the proof is to show that the continuous functional calculus can be represented as integration against a projection-valued measure. This result is just an operator-valued version of the Riesz representation theorem from measure theory (Theorem 8.5). Indeed, we will see that this operator-valued version of the Riesz representation theorem can be reduced to the usual form of the theorem.

8.1.1 Stage 1: The Continuous Functional Calculus

We begin by defining, for any $A \in \mathcal{B}(\mathbf{H})$, the *spectral radius* $R(A)$ by

$$R(A) = \sup_{\lambda \in \sigma(A)} |\lambda|.$$

(By Propositions 7.5 and 7.7, $\sigma(A)$ is a nonempty, bounded subset of \mathbb{R} .) According to Point 2 of Proposition 7.5, we have

$$R(A) \leq \|A\|$$

for any $A \in \mathcal{B}(\mathbf{H})$. In general, $\|A\|$ can be much bigger than $R(A)$. For example, if A is a nilpotent matrix, then $R(A) = 0$ but $\|A\|$ can be arbitrarily large.

Lemma 8.1 *If $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint, the norm and the spectral radius of A are equal:*

$$\|A\| = R(A).$$

In preparation for the proof, we determine the radius of convergence of the power series for the resolvent given in the proof of Proposition 7.5. According to Proposition 7.2, we have

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

for any $A \in \mathcal{B}(\mathbf{H})$. If A is self-adjoint, we obtain

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

Iterating this relation gives

$$\|A^{2^n}\| = \|A\|^{2^n} \tag{8.1}$$

for all n .

Consider, for a bounded self-adjoint operator A , the following formal expression for the resolvent of A :

$$\begin{aligned} (A - \lambda I)^{-1} &= -\frac{1}{\lambda} \left(I - \frac{A}{\lambda} \right)^{-1} \\ &= -\sum_{m=0}^{\infty} \frac{A^m}{\lambda^{m+1}}. \end{aligned} \tag{8.2}$$

If $|\lambda| > \|A\|$, then the proof of Proposition 7.5 shows that the series (8.2) converges in the operator norm topology and that the sum of the series is indeed the inverse of $(A - \lambda I)$. If, on the other hand, $|\lambda| \leq \|A\|$, it follows from (8.1) that the norms of the terms in (8.2) do not tend to zero, and

so the series cannot converge in the operator norm topology. We may say, then, that the series (8.2) has radius of convergence equal to $\|A\|$.

Proof of Lemma 8.1. We know that $R(A) \leq \|A\|$. To show that $R(A) = \|A\|$, we wish to argue that $(A - \lambda I)^{-1}$ is a holomorphic operator-valued function of λ on the set $|\lambda| > R(A)$, and therefore the Laurent series of $(A - \lambda I)^{-1}$ must converge for $|\lambda| > R(A)$. But the Laurent series of $(A - \lambda I)^{-1}$ is just the series in (8.2), and we have shown that the series diverges when $|\lambda| \leq \|A\|$. This would be a contradiction if $R(A)$ were less than $\|A\|$.

To flesh out the argument, recall the formula (7.8) in the proof of Proposition 7.5 for the resolvent of A .

That formula expresses the map $\lambda \mapsto (A - \lambda I)^{-1}$ as a convergent power series in powers of $\lambda - \lambda_0$, near any point λ_0 in the resolvent set of A . It follows that for any bounded linear functional $\xi \in \mathcal{B}(\mathbf{H})^*$, the complex-valued function

$$\lambda \mapsto \xi((A - \lambda I)^{-1})$$

is holomorphic on the resolvent set of A . This function has a unique Laurent series, which is given by applying ξ term by term to (8.2). The series will converge on the largest annulus contained in the resolvent set of A , namely the set of λ with $|\lambda| > R(A)$.

Convergence of (8.2) means that $|\xi(A^m/\lambda^{m+1})|$ is bounded as function of m , for each ξ and each λ with $|\lambda| > R(A)$. Thus, by (a corollary of) the uniform boundedness principle (Appendix A.3.4), the set $\{A^m/\lambda^{m+1}\}_{m=0}^{\infty}$ is bounded in the Banach space $\mathcal{B}(\mathbf{H})$, for all λ with $|\lambda| > R(A)$. In particular, for each λ with $|\lambda| > R(A)$, there is a constant C such that

$$\frac{\|A^{2^n}\|}{|\lambda|^{2^n}} = \frac{\|A\|^{2^n}}{|\lambda|^{2^n}} \leq C.$$

If $\|A\|$ were greater than $R(A)$, this inequality would be false for λ satisfying $R(A) < |\lambda| < \|A\|$. ■

The next key step in Stage 1 of the proof is to understand how the spectrum of a self-adjoint operator transforms under application of a polynomial.

Lemma 8.2 (Spectral Mapping Theorem) *For all $A \in \mathcal{B}(\mathbf{H})$ and all polynomials p , we have*

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

That is to say, the spectrum of $p(A)$ consists precisely of the numbers of the form $p(\lambda)$, with λ in the spectrum of A .

Proof. The result is trivial if p is constant. When $\deg p \geq 1$, let p given by

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

be an arbitrary polynomial. We first show that $p(\sigma(A)) \subset \sigma(p(A))$. Suppose, then, that $\lambda \in \sigma(A)$. Observe that

$$p(A) - p(\lambda)I = a_n(A^n - \lambda^n I) + a_{n-1}(A^{n-1} - \lambda^{n-1}I) + \cdots + a_0I - a_0I.$$

Now,

$$A^k - \lambda^k I = (A - \lambda I)(A^{k-1} + \lambda A^{k-2} + \lambda^2 A^{k-3} + \cdots + \lambda^{k-1}I).$$

Thus, we can pull out a factor of $(A - \lambda I)$ from each nonzero term in $p(A) - p(\lambda)I$, giving

$$p(A) - p(\lambda)I = (A - \lambda I)q(A)$$

where q is a polynomial (depending on λ). Since, by assumption, $A - \lambda I$ is not invertible, and since $(A - \lambda I)$ commutes with $q(A)$, $(A - \lambda I)q(A)$ cannot be invertible (Exercise 1). This shows that $p(\lambda)$ belongs to the spectrum of $p(A)$.

We now show that $\sigma(p(A)) \subset p(\sigma(A))$. Suppose, then, that $\gamma \in \sigma(p(A))$. Since \mathbb{C} is algebraically closed, we can factor the polynomial $p(z) - \gamma$, as a function of z , as

$$p(z) - \gamma = c(z - b_1)(z - b_2) \cdots (z - b_n). \quad (8.3)$$

Thus,

$$p(A) - \gamma I = c(A - b_1 I)(A - b_2 I) \cdots (A - b_n I).$$

Since $p(A) - \gamma I$ is assumed to be noninvertible, there must be some j such that $(A - b_j I)$ is noninvertible, that is, for which $b_j \in \sigma(A)$. Then (8.3) tells us that $p(b_j) - \gamma = 0$, meaning that $\gamma = p(b_j)$. Thus, γ is of the form $p(\lambda)$ for some $\lambda (= b_j)$ in $\sigma(A)$. ■

The last step in Stage 1 of our proof is to apply the Stone–Weierstrass theorem to show that polynomials are dense in $\mathcal{C}(\sigma(A); \mathbb{R})$ (the space of continuous, real-valued functions on $\sigma(A)$) with respect to the supremum norm.

Proposition 8.3 *Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint. Then there exists a unique bounded linear map from $\mathcal{C}(\sigma(A); \mathbb{R})$ into $\mathcal{B}(\mathbf{H})$, denoted by $f \mapsto f(A)$, such that when $f(\lambda) = \lambda^m$, we have $f(A) = A^m$. The map $f \mapsto f(A)$, $f \in \mathcal{C}(\sigma(A); \mathbb{R})$, is called the (real-valued) functional calculus for A .*

Proof. Note that if A is self-adjoint, then $p(A)$ is self-adjoint provided that p is a *real-valued* polynomial (i.e., one where all the coefficients are real numbers). Thus, combining the spectral mapping theorem with the equality of the norm and spectral radius, we have the following: If A is a self-adjoint operator and p is a real-valued polynomial, then

$$\|p(A)\| = \sup_{\lambda \in \sigma(A)} |p(\lambda)|. \quad (8.4)$$

Thus, the map $p \rightarrow p(A)$ is an isometric linear map from the space of polynomials on $\sigma(A)$ (with the supremum norm) into the space of bounded operators on \mathbf{H} .

According to the Stone–Weierstrass theorem polynomials are dense in $\mathcal{C}(\sigma(A); \mathbb{R})$. Thus, by the BLT theorem (Theorem A.36), we can extend the map $p \mapsto p(A)$ uniquely to a bounded linear map of $\mathcal{C}(\sigma(A); \mathbb{R})$ into $\mathcal{B}(\mathbf{H})$.

■

Proposition 8.4 *If $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint, the (real-valued) continuous functional calculus for A , mapping $\mathcal{C}(\sigma(A); \mathbb{R})$ into $\mathcal{B}(\mathbf{H})$, has the following properties.*

1. **Multiplicativity:** For all f, g , we have

$$(fg)(A) = f(A)g(A),$$

where fg denotes the pointwise product of f and g .

2. **Self-adjointness:** For all f , the operator $f(A)$ is self-adjoint.

3. **Non-negativity:** For all f , if f is non-negative, then $f(A)$ is a non-negative operator.

4. **Norm and spectrum properties:** For all f , we have

$$\|f(A)\| = \sup_{\lambda \in \sigma(A)} |f(\lambda)| \quad (8.5)$$

and

$$\sigma(f(A)) = \{f(\lambda) \mid \lambda \in \sigma(A)\}. \quad (8.6)$$

Proof. Point 1 holds for polynomials and thus, by taking limits, for all $f \in \mathcal{C}(\sigma(A); \mathbb{R})$. Furthermore, if p is a real-valued polynomial and A is self-adjoint, then $p(A)$ is self-adjoint. From this, we get Point 2 by taking limits. If $f \in \mathcal{C}(\sigma(A); \mathbb{R})$ is non-negative, then $f = g^2$, where $g = \sqrt{f}$ is real-valued. Thus, $g(A)$ is self-adjoint and for all $\psi \in \mathbf{H}$, Point 1 tells us that

$$\langle \psi, f(A)\psi \rangle = \langle \psi, g(A)^2\psi \rangle = \langle g(A)\psi, g(A)\psi \rangle \geq 0, \quad (8.7)$$

which establishes Point 3. We have already established (8.5) in (8.4) for polynomials; the result for general $f \in \mathcal{C}(\sigma(A); \mathbb{R})$ follows by taking limits.

To establish (8.6), suppose first that $\lambda_0 \in \mathbb{C}$ is not in the range of f . Then the function $g(\lambda) := 1/(f(\lambda) - \lambda_0)$ is continuous on $\sigma(A)$ and the operator $g(A)$ will be the inverse of $f(A) - \lambda_0 I$, showing that λ_0 is not in the spectrum of $f(A)$.

In the other direction, suppose that $\lambda_0 = f(\mu)$ for some $\mu \in \sigma(A)$; we want to show that $f(\mu) \in \sigma(f(A))$. Suppose now that $f(A) - f(\mu)I$ were invertible and choose a sequence p_n of polynomials converging uniformly

to f on $\sigma(A)$. By Exercise 8 in Chap. 7, any operator sufficiently close to $f(A) - f(\mu)I$ in the operator norm topology would also be invertible. In particular, $p_n(A) - p_n(\mu)I$ would have to be invertible for all sufficiently large n , contradicting the spectral mapping theorem. ■

8.1.2 Stage 2: An Operator-Valued Riesz Representation Theorem

We turn now to Stage 2 of the proof of the spectral theorem. We will make use of the Riesz representation theorem from measure theory (*not* the result about continuous linear functionals on a Hilbert space). The following form of this result is sufficient for our purposes.

Theorem 8.5 (Riesz Representation Theorem) *Let X be a compact metric space and let $\mathcal{C}(X; \mathbb{R})$ denote the space of continuous, real-valued functions on X . Suppose $\Lambda : \mathcal{C}(X; \mathbb{R}) \rightarrow \mathbb{R}$ is a linear functional with the property that $\Lambda(f)$ is non-negative whenever all the values of f are non-negative. Then there exists a unique (real-valued, positive) measure μ on the Borel σ -algebra in X for which*

$$\Lambda(f) = \int_X f \, d\mu$$

for all $f \in \mathcal{C}(X; \mathbb{R})$.

See pp. 353–354 of Volume I of [34] for a short proof in the case in which X is a compact subset of \mathbb{R} , which is all we really require. For the full result stated above, see Theorems 7.2 and 7.8 in [12]. Observe that μ is a finite measure, with $\mu(X) = \Lambda(\mathbf{1})$, where $\mathbf{1}$ is the constant function.

Given a bounded self-adjoint operator $A \in \mathcal{B}(\mathbf{H})$, we have constructed, in the previous subsection, a continuous functional calculus for A . This calculus is a map, denoted $f \mapsto f(A)$, from $\mathcal{C}(\sigma(A); \mathbb{R})$ into $\mathcal{B}(\mathbf{H})$. If $f \in \mathcal{C}(\sigma(A); \mathbb{R})$ is non-negative, then (Point 3 of Proposition 8.4) $f(A)$ is a non-negative operator. Thus, given $\psi \in \mathbf{H}$, if we define a linear functional Λ_ψ on $\mathcal{C}(\sigma(A); \mathbb{R})$ by the formula

$$\Lambda_\psi(f) = \langle \psi, f(A)\psi \rangle,$$

Λ_ψ will satisfy the hypotheses of the Riesz representation theorem. Thus, for each $\psi \in \mathbf{H}$, we obtain a unique measure μ_ψ such that

$$\langle \psi, f(A)\psi \rangle = \int_{\sigma(A)} f(\lambda) \, d\mu_\psi(\lambda) \tag{8.8}$$

for all $f \in \mathcal{C}(\sigma(A); \mathbb{R})$. Note that

$$\mu_\psi(\sigma(A)) = \Lambda_\psi(\mathbf{1}) = \|\psi\|^2. \tag{8.9}$$

Definition 8.6 If f is a bounded measurable (complex-valued) function on $\sigma(A)$, define a map $Q_f : \mathbf{H} \rightarrow \mathbb{C}$ by the formula

$$Q_f(\psi) = \int_{\sigma(A)} f(\lambda) d\mu_\psi(\lambda),$$

where μ_ψ is the measure in (8.8).

If f happens to be real valued and continuous, then $Q_f(\psi)$ is equal $\langle \psi, f(A)\psi \rangle$, in which case Q_f is a bounded quadratic form. (See Definition A.60 and Example A.62.) It turns out that Q_f is a bounded quadratic form for any bounded measurable f , in which case Proposition A.63 allows us to associate with Q_f a bounded operator, which we denote by $f(A)$. Once the relevant properties of $f(A)$ are established, we will construct the desired projection-valued measure by setting $\mu^A(E) = 1_E(A)$.

Proposition 8.7 For any bounded measurable function f on $\sigma(A)$, the map Q_f in Definition 8.6 is a bounded quadratic form.

Proof. Let \mathcal{F} denote the space of all bounded, Borel-measurable functions f for which Q_f is a quadratic form. Then \mathcal{F} is a vector space and contains $\mathcal{C}(\sigma(A); \mathbb{R})$. Furthermore, \mathcal{F} is closed under uniformly bounded pointwise limits, because $Q_f(\psi)$ is continuous with respect to such limits, by dominated convergence. Standard measure-theoretic techniques (Exercise 3) then show that \mathcal{F} is the space of all bounded Borel-measurable functions on X .

Meanwhile, it follows from (8.9) that

$$|Q_f(\psi)| \leq \sup_{\lambda \in \sigma(A)} |f(\lambda)| \|\psi\|^2,$$

showing that Q_f is always a bounded quadratic form. ■

Definition 8.8 For a bounded measurable function f on $\sigma(A)$, let $f(A)$ be the operator associated to the quadratic form Q_f by Proposition A.63. This means that $f(A)$ is the unique operator such that

$$\langle \psi, f(A)\psi \rangle = Q_f(\psi) = \int_{\sigma(A)} f d\mu_\psi$$

for all $\psi \in \mathbf{H}$.

Observe that if f is real valued, then $Q_f(\psi)$ is real for all $\psi \in \mathbf{H}$, which means (Proposition A.63) that the associated operator $f(A)$ is self-adjoint. We will shortly associate with A a projection-valued measure μ^A , and we will show that $f(A)$, as given by Definition 8.8, agrees with $f(A)$ as given by $\int_{\sigma(A)} f(\lambda) d\mu^A(\lambda)$. [See (8.10) and compare Definition 7.13.]

Proposition 8.9 *For any two bounded measurable functions f and g , we have*

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

Proof. Let \mathcal{F}_1 denote the space of bounded measurable functions f such that $(fg)(A) = f(A)g(A)$ for all $g \in \mathcal{C}(\sigma(A); \mathbb{R})$. Then \mathcal{F}_1 is a vector space and contains $\mathcal{C}(\sigma(A); \mathbb{R})$. We have already noted that dominated convergence guarantees that the map $f \mapsto Q_f(\psi)$, $\psi \in \mathbf{H}$, is continuous under uniformly bounded pointwise convergence. By the polarization identity (Proposition A.59), the same is true for the map $f \mapsto L_f(\phi, \psi)$, where L_f is the sesquilinear form associated to Q_f . Now, by the polarization identity, f will be in \mathcal{F}_1 provided that

$$\langle \psi, (fg)(A)\psi \rangle = \langle \psi, f(A)g(A)\psi \rangle$$

or, equivalently,

$$Q_{fg}(\psi) = L_f(\psi, g(A)\psi)$$

for all $\psi \in \mathbf{H}$ and all $g \in \mathcal{C}(\sigma(A); \mathbb{R})$. From this, we can see that \mathcal{F}_1 is closed under uniformly bounded pointwise limits. Thus, by Exercise 3, \mathcal{F}_1 consists of all bounded, Borel-measurable functions.

We now let \mathcal{F}_2 denote the space of all bounded, Borel-measurable functions f such that $(fg)(A) = f(A)g(A)$ for all bounded Borel-measurable functions g . Our result for \mathcal{F}_1 shows that \mathcal{F}_2 contains $\mathcal{C}(\sigma(A); \mathbb{R})$. Thus, the same argument as for \mathcal{F}_1 shows that \mathcal{F}_2 consists of all bounded, Borel-measurable functions. ■

Theorem 8.10 *Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint. For any measurable set $E \subset \sigma(A)$, define an operator $\mu^A(E)$ by*

$$\mu^A(E) = 1_E(A),$$

where $1_E(A)$ is given by Definition 8.8. Then μ^A is a projection-valued measure on $\sigma(A)$ and satisfies

$$\int_{\sigma(A)} \lambda d\mu^A(\lambda) = A.$$

Theorem 8.10 establishes the existence of the projection-valued measure in our first version of the spectral theorem (Theorem 7.12).

Proof. Since 1_E is real-valued and satisfies $1_E \cdot 1_E = 1_E$, Proposition 8.4 tells us that $1_E(A)$ is self-adjoint and satisfies $1_E(A)^2 = 1_E(A)$. Thus, $\mu^A(E)$ is an orthogonal projection (Proposition A.57), for any measurable set $E \subset X$. If E_1 and E_2 are measurable sets, then $1_{E_1 \cap E_2} = 1_{E_1} \cdot 1_{E_2}$ and so

$$\mu^A(E_1 \cap E_2) = \mu^A(E_1)\mu^A(E_2).$$

If E_1, E_2, \dots are disjoint measurable sets, then $\mu^A(E_j)\mu^A(E_k) = \mu^A(\emptyset) = 0$, for $j \neq k$, and so the ranges of the projections $\mu^A(E_j)$ and $\mu^A(E_k)$ are

orthogonal. It then follows by an elementary argument that, for all $\psi \in \mathbf{H}$, we have

$$\sum_{j=1}^{\infty} \mu^A(E_j)\psi = P\psi,$$

where the sum converges in the norm topology of \mathbf{H} and where P is the orthogonal projection onto the smallest closed subspace containing the range of $\mu^A(E_j)$ for every j . On the other hand, if $E := \cup_{j=1}^{\infty} E_j$, then the sequence $f_N := \sum_{j=1}^N 1_{E_j}$ is uniformly bounded (by 1) and converges pointwise to 1_E . Thus, using again dominated convergence in (8.8),

$$\lim_{N \rightarrow \infty} \left\langle \psi, \sum_{j=1}^N 1_{E_j}(A)\psi \right\rangle = \langle \psi, 1_E(A)\psi \rangle.$$

It follows that $1_E(A)$ coincides with P , which establishes the desired countable additivity for μ^A .

Finally, if $f = 1_E$ for some Borel set E , then

$$\int_{\sigma(A)} f(\lambda) d\mu^A(\lambda) = f(A), \quad (8.10)$$

where $f(A)$ is given by Definition 8.8. [The integral is equal to $\mu^A(E)$, which is, by definition, equal to $1_E(A)$.] The equality (8.10) then holds for simple functions by linearity and for all bounded, Borel-measurable functions by taking limits. In particular, if $f(\lambda) = \lambda$, then the integral of f against μ^A agrees with $f(A)$ as defined in Definition 8.8, which agrees with $f(A)$ as defined in the continuous functional calculus, which in turn agrees with $f(A)$ as defined for polynomials—namely, $f(A) = A$. This means that

$$\int_{\sigma(A)} \lambda d\mu^A(\lambda) = A$$

as desired. ■

We have now completed the existence of the projection-valued measure μ^A in Theorem 7.12. The uniqueness of μ^A is left as an exercise (Exercise 4). We close this section by proving Proposition 7.16, which states that if a bounded operator B commutes with a bounded self-adjoint operator A , then B commutes with $f(A)$, for all bounded, Borel-measurable functions f on $\sigma(A)$.

Proof of Proposition 7.16. If B commutes with A , then B commutes with $p(A)$, for any polynomial p . Thus, by taking limits as in the construction of the continuous functional calculus, B will commute with $f(A)$ for any continuous real-valued function f on $\sigma(A)$. We now let \mathcal{F} denote the space of all bounded, Borel-measurable functions f on $\sigma(A)$ for which $f(A)$ commutes with B , so that $\mathcal{C}(\sigma(A); \mathbb{R})$.

To show that a bounded measurable f belongs to \mathcal{F} , it suffices to show that for all $\phi, \psi \in \mathbf{H}$ we have $\langle \phi, f(A)B\psi \rangle = \langle \phi, Bf(A)\psi \rangle$, or, equivalently, $\langle \phi, f(A)B\psi \rangle = \langle B^*\phi, f(A)\psi \rangle$. That is, we want

$$L_f(\phi, B\psi) = L_f(B^*\phi, \psi).$$

But we have seen that for fixed vectors $\psi_1, \psi_2 \in \mathbf{H}$, the map $f \mapsto L_f(\psi_1, \psi_2)$ is continuous under uniformly bounded pointwise limits. Thus, \mathcal{F} is closed under such limits, which implies (Exercise 3) that \mathcal{F} contains all bounded, Borel-measurable functions. ■

8.2 Proof of the Spectral Theorem, Second Version

We now turn to the proof of Theorem 7.19. As in the proof of Theorem 7.12, we will make use of continuous functional calculus for a bounded self-adjoint operator A and the Riesz representation theorem. We begin by establishing the special case in which A has a **cyclic vector**, that is, a vector ψ with the property that the vectors $A^k\psi$, $k = 0, 1, 2, \dots$, span a dense subspace of \mathbf{H} . In that case, the direct integral will be simply an L^2 space (i.e., the Hilbert spaces \mathbf{H}_λ are equal to \mathbb{C} for all λ). Thus, in this special case, the direct integral and multiplication operator versions of the spectral theorem coincide.

Lemma 8.11 *Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint and ψ is a cyclic vector for A . Let μ_ψ be the unique measure on $\sigma(A)$, given by Theorem 8.5, for which*

$$\langle \psi, f(A)\psi \rangle = \int_{\sigma(A)} f(\lambda) d\mu_\psi(\lambda) \quad (8.11)$$

for all $f \in \mathcal{C}(\sigma(A); \mathbb{R})$. Then there exists a unitary map

$$U : \mathbf{H} \rightarrow L^2(\sigma(A), \mu_\psi)$$

such that

$$[UAU^{-1}\phi](\lambda) = \lambda\phi(\lambda)$$

for all $\phi \in L^2(\sigma(A), \mu_\psi)$.

Proof. We start by defining U on the complex vector space of vectors of the form $p(A)\psi$, where p is a complex-valued polynomial, as follows:

$$U[p(A)\psi] = p.$$

To show that U is well defined, write p as $p = p_1 + ip_2$, where p_1 and p_2 are real-valued polynomials. Since $p_1(A)$ and $p_2(A)$ are self-adjoint and

commuting, we obtain

$$\begin{aligned} \langle p(A)\psi, p(A)\psi \rangle &= \langle \psi, [p_1(A)^2 + p_2(A)^2] \psi \rangle \\ &= \int_{\sigma(A)} [p_1(\lambda)^2 + p_2(\lambda)^2] d\mu_\psi(\lambda), \end{aligned} \quad (8.12)$$

by canceling cross terms and applying (8.11). Thus, if $p(A)\psi = 0$ in \mathbf{H} , then $p(\lambda) = 0$ for μ_ψ -almost every λ in $\sigma(A)$, so that p represents the zero element of $L^2(\sigma(A), \mu_\psi)$.

Equation (8.12) shows also that the map U is isometric on its initial domain. This initial domain is dense in \mathbf{H} since it contains the vectors $A^k\psi$ and ψ is cyclic. Thus, the BLT theorem (Theorem A.36) tells us that U extends uniquely to an isometric map of \mathbf{H} into $L^2(\sigma(A), \mu_\psi)$. Since polynomials are dense in $L^2(\sigma(A), \mu_\psi)$ (by the Stone–Weierstrass theorem and Theorem A.10), U actually is unitary.

Now, since U takes $A^k\psi$ to the function $\lambda \mapsto \lambda^k$ in $L^2(\sigma(A), \mu_\psi)$, we have that $UAU^{-1}(\lambda^k) = \lambda^{k+1}$. Thus,

$$[UAU^{-1}p](\lambda) = \lambda p(\lambda)$$

for all polynomials p . Since polynomials are dense in $L^2(\sigma(A), \mu_\psi)$, we have $[UAU^{-1}\phi](\lambda) = \lambda\phi(\lambda)$ for all $\phi \in L^2(\sigma(A), \mu_\psi)$, as claimed. ■

Lemma 8.12 *Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint and μ^A is the associated projection-valued measure on $\sigma(A)$, as in Theorem 8.10. Then there exists a non-negative real-valued measure μ on $\sigma(A)$ such that for all Borel sets $E \subset \sigma(A)$, we have $\mu^A(E) = 0$ if and only if $\mu(E) = 0$.*

Proof. Let $\{e_j\}$ be an orthonormal basis for \mathbf{H} and let μ_{e_j} be the associated real-valued measures, given by $\mu_{e_j}(E) = \langle e_j, \mu^A(E)e_j \rangle$. Then $\mu_{e_j}(\sigma(A)) = \langle e_j, Ie_j \rangle = 1$ for all j . Thus, the formula

$$\mu := \sum_j \frac{1}{j^2} \mu_{e_j}$$

defines a finite measure on $\sigma(A)$. Given some Borel set $E \subset \sigma(A)$, if $\mu^A(E) = 0$, then $\mu_{e_j}(E) = 0$ for all j and so $\mu(E) = 0$. Conversely, if $\mu(E) = 0$, then

$$0 = \langle e_j, \mu^A(E)e_j \rangle = \langle \mu^A(E)e_j, \mu^A(E)e_j \rangle$$

for all j , since $\mu^A(E)$ is self-adjoint and $\mu^A(E)^2 = \mu^A(E)$. Thus, $\mu^A(E)e_j = 0$ for all j , which means that $\mu^A(E) = 0$. ■

Lemma 8.13 *If $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint, then \mathbf{H} can be decomposed as an orthogonal direct sum of closed nonzero subspaces W_j , where each W_j is invariant under A and where the restriction of A to W_j has a cyclic vector ψ_j . The number of W_j 's is either finite or countably infinite.*

Proof. Recall our standing assumption that \mathbf{H} is separable, and let $\{\phi_j\}$ be a countable dense subset of \mathbf{H} . Let W_1 be the closed subspace of \mathbf{H} spanned by $\phi_1, A\phi_1, A^2\phi_1, \dots$. Then W_1 is invariant under A and $\psi_1 := \phi_1$ is a cyclic vector for $A|_{W_1}$. If $W_1 = \mathbf{H}$ then we are done. If not, let j be the smallest number such that ϕ_j is not contained in W_1 . Let ψ_2 be the orthogonal projection of ϕ_j onto the orthogonal complement of W_1 , and let W_2 be the closed span of $\psi_2, A\psi_2, A^2\psi_2, \dots$. Then W_2 is invariant under A and ψ_2 is a cyclic vector for $A|_{W_2}$. Furthermore, since A is self-adjoint and leaves W_1 invariant, it also leaves W_1^\perp invariant, which means that $A^k\psi_2$ is orthogonal to W_1 for all k , so that W_2 is orthogonal to W_1 .

If, now, $W_1 \oplus W_2 = \mathbf{H}$, we are done. If not, we let k be the smallest number such that ϕ_k is not in $W_1 \oplus W_2$ and we let ψ_3 be the projection of ϕ_k onto the orthogonal complement of $W_1 \oplus W_2$, and so on. Continuing on in this way, we obtain an orthogonal collection of closed subspaces that are invariant under A , each of which has a cyclic vector. Either the process terminates with finitely many of these subspaces spanning \mathbf{H} , or we get an infinite family. In the latter case, each ϕ_j belongs to the span of the W_j 's and hence the (Hilbert space) direct sum of the W_j 's is all of \mathbf{H} . ■

We are now ready for the proof of our second form of the spectral theorem.

Proof of Theorem 7.19. Let $\{W_j, \psi_j\}$ be as in Lemma 8.13, and let A_j denote the restriction of A to W_j , which is a bounded self-adjoint operator on the Hilbert space W_j . For each A_j , we can obtain a unitary map U_j as in Lemma 8.11, and we wish to piece these maps together for different values of j to obtain a direct integral decomposition for all of \mathbf{H} . To facilitate piecing the maps together, we will modify the U_j 's so that they all map to L^2 spaces over a subset of $\sigma(A)$ with respect to the *same measure* μ .

If we apply Lemma 8.11 to A_j , we get a unitary map

$$U_j : W_j \rightarrow L^2(\sigma(A_j), \mu_{\psi_j})$$

such that $U_j A U_j^{-1}$ is the operator of multiplication by λ . Here, μ_{ψ_j} is the measure on $\sigma(A_j)$ given by $\mu_{\psi_j}(E) = \langle \psi_j, \mu^{A_j}(E) \psi_j \rangle$. Now, according to Exercise 5, the spectrum of A_j is contained in the spectrum of A . Furthermore, if E is a measurable subset of $\sigma(A_j) \subset \sigma(A)$, then 1_E may be thought of as a measurable function either on $\sigma(A_j)$ or on $\sigma(A)$. Exercise 5 tells us that $1_E(A_j)$, as defined by the functional calculus for A_j , coincides with the restriction to W_j of $1_E(A)$. Thus, if $1_E(A) = 0$ then $1_E(A_j) = 0$ as well. Equivalently, if $\mu^A(E) = 0$ then $\mu^{A_j}(E) = 0$, where μ^{A_j} is the projection-valued measure associated to the self-adjoint operator A_j .

Let us now choose a measure μ as in Lemma 8.12. Any set of measure zero for μ is a set of measure zero for μ^A and thus also for μ^{A_j} and then for μ_{ψ_j} . Thus, if we extend μ_{ψ_j} to a measure on $\sigma(A)$ by making it zero on $\sigma(A) \setminus \sigma(A_j)$, we have that μ_{ψ_j} is absolutely continuous with respect to μ .

By the Radon–Nikodym theorem (Theorem A.6), each μ_{ψ_j} has a density ρ_j with respect to μ , and this density is nonzero μ_{ψ_j} -almost everywhere.

Now, the map

$$f \mapsto \rho_j^{1/2} f$$

is easily seen to be a unitary map of $L^2(\sigma(A_j), \mu_{\psi_j})$ to $L^2(\sigma(A_j), \mu)$. Thus, we can define a unitary map

$$\tilde{U}_j : W_j \rightarrow L^2(\sigma(A_j), \mu)$$

by setting

$$(\tilde{U}_j \psi)(\lambda) = \rho_j(\lambda)^{1/2} (U_j \psi)(\lambda).$$

Since multiplication by $(\rho_j)^{1/2}$ commutes with multiplication by λ , we have

$$\left(\tilde{U}_j A_j \tilde{U}_j^{-1} \right) (\psi)(\lambda) = \lambda \psi(\lambda).$$

Now, $L^2(\sigma(A_j), \mu)$ can be thought of as a direct integral over $\sigma(A)$ with respect to μ , where we take $\mathbf{H}_\lambda^j = \mathbb{C}$ for $\lambda \in \sigma(A_j)$ and we take $\mathbf{H}_\lambda^j = \{0\}$ if $\lambda \in \sigma(A_j)^c$. We now define another direct integral over $\sigma(A)$ in which the Hilbert spaces \mathbf{H}_λ , $\lambda \in \sigma(A)$, are defined by

$$\mathbf{H}_\lambda = \bigoplus_j \mathbf{H}_\lambda^j.$$

Here the measurable structure on the direct integral is defined by setting

$$e_j(\lambda) = \begin{cases} (0, 0, \dots, 1, 0, 0, \dots), & \lambda \in E_j \\ (0, 0, \dots, 0, 0, 0, \dots), & \lambda \in E_j^c \end{cases},$$

where the 1 is in the j th slot. Since each \mathbf{H}_λ is a direct sum of the \mathbf{H}_λ^j 's, the direct integral of the \mathbf{H}_λ 's is the Hilbert space direct sum of the direct integral of the \mathbf{H}_λ^j 's, which is just $L^2(\sigma(A_j), \mu)$.

Meanwhile, \mathbf{H} is the direct sum of the W_j 's, and we have unitary maps \tilde{U}_j of W_j to $L^2(\sigma(A_j), \mu)$ such that $\tilde{U}_j A_j \tilde{U}_j^{-1}$ is just multiplication by λ on $L^2(E_j, \mu)$. Thus, we can assemble the \tilde{U}_j 's into a single unitary map U of \mathbf{H} to the integral of the \mathbf{H}_λ 's, and we will have $U A U^{-1}$ equal to multiplication by λ , as desired. ■

In the interest of brevity, we will not give a complete proof of Proposition 7.22 (uniqueness in Theorem 7.19), but only indicate the main ideas. To establish the equivalence of $\mu^{(1)}$ and $\mu^{(2)}$, we observe that both measures have the same sets of measure zero as the projection-valued measure μ^A (Proposition 7.23). Meanwhile, if we have two different direct integrals, each unitarily equivalent to \mathbf{H} as in (7.20), then there will be a unitary map V between the two direct integrals that commutes with the operator $s(\lambda) \mapsto \lambda s(\lambda)$. Using an argument similar to that in Exercise 7, we

can show that there must be bounded maps $V_\lambda : \mathbf{H}_\lambda^{(1)} \rightarrow \mathbf{H}_\lambda^{(2)}$ such that $(Vs)(\lambda) = V_\lambda s(\lambda)$ for almost every λ . Then we argue that the only way V can be unitary is if V_λ is unitary for almost every λ . This implies that $\dim \mathbf{H}_\lambda^{(1)} = \dim \mathbf{H}_\lambda^{(2)}$ for almost every λ .

Finally, we briefly indicate the proof of the multiplication operator form of the spectral theorem.

Proof of Theorem 7.20. Let W_j be as in Lemma 8.13 and let A_j be the restriction of A to W_j . By the proof of Theorem 7.19, each A_j is unitarily equivalent to multiplication by λ on the Hilbert space $L^2(\sigma(A_j), \mu_j)$, for some finite measure μ_j on $\sigma(A_j)$. Let X be the disjoint union of the sets $\sigma(A_j)$, let μ be the sum of the measures μ_j , and let h be the function whose restriction to each $\sigma(A_j)$ is the function $\lambda \mapsto \lambda$. Then $L^2(X, \mu)$ is the orthogonal direct sum of the Hilbert spaces $L^2(\sigma(A_j), \mu_j)$, which means that $L^2(X, \mu)$ may be identified unitarily with $\mathbf{H} = \oplus W_j$ in an obvious way. Under this identification, the operator A corresponds to multiplication by h .

■

8.3 Exercises

1. (a) Suppose $A, B \in \mathcal{B}(\mathbf{H})$ commute and A is not invertible. Show that AB is not invertible.
Hint: First show that if AB were invertible, then A would have both a left inverse and a right inverse. Then show that the left inverse and right inverse would need to be equal.
- (b) Show that the result of Part (a) is false if we omit the assumption that A and B commute.
2. (a) Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint and $\sigma(A) \subset [0, \infty)$. Show that A has a self-adjoint square root in $\mathcal{B}(\mathbf{H})$ and therefore that A is a non-negative operator (i.e., $\langle \psi, A\psi \rangle \geq 0$ for all $\psi \in \mathbf{H}$).
- (b) Give an example of a bounded operator A on a Hilbert space such that $\sigma(A) \subset [0, \infty)$ but A is not non-negative.
3. Let X be a compact metric space and let $\mathcal{C}(X; \mathbb{R})$ denote the space of continuous real-valued functions on X . Suppose that \mathcal{F} is a set of bounded, measurable, complex-valued functions on X with the following properties: (1) \mathcal{F} is a complex vector space, (2) \mathcal{F} contains $\mathcal{C}(X; \mathbb{R})$, and (3) \mathcal{F} is closed under pointwise limits of uniformly bounded sequences. (A sequence f_n is uniformly bounded if there exists a constant C such that $|f_n(x)| \leq C$ for all n and x).
- (a) Let \mathcal{L}_0 denote the collection of those measurable sets E for which $\mathbf{1}_E$ is a uniformly bounded limit of a sequence of continuous

functions. Show that \mathcal{L}_0 is an algebra and contains all open sets in X .

(b) Let \mathcal{L}_1 denote the collection of all measurable sets in E for which 1_E belongs to \mathcal{F} . Using the monotone class lemma (Theorem A.8), show that \mathcal{L}_1 consists of all Borel sets in X .

(c) Show that \mathcal{F} consists of all bounded, Borel-measurable functions on X .

4. Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint μ^A and ν^A are two projection-valued measures on $\sigma(A)$ such that

$$\int_{\sigma(A)} \lambda d\mu^A(\lambda) = \int_{\sigma(A)} \lambda d\nu^A(\lambda) = A.$$

Show that integration with respect to μ^A agrees with integration with respect to ν^A , first on polynomials, then on continuous functions, and finally on bounded measurable functions. Conclude that $\mu^A = \nu^A$.

Hint: Use Exercise 17.

5. Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint operator and V is a closed subspace of \mathbf{H} that is invariant under A .

(a) Using Proposition 7.7, show that the spectrum of the restriction to V of A is contained in the spectrum of A .

(b) Suppose now that f is a bounded measurable function on $\sigma(A)$, which means that f is also a function on $\sigma(A|_V) \subset \sigma(A)$. Show that V is invariant under $f(A)$ and that

$$f(A)|_V = f(A|_V),$$

where the operator on the right-hand side is defined by the measurable functional calculus for the bounded self-adjoint operator $A|_V$.

6. Suppose $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint and ψ is an eigenvector for A , that is, a nonzero vector with $A\psi = \lambda\psi$ for some $\lambda \in \mathbb{R}$. Show that for any bounded measurable function f on $\sigma(A)$ we have

$$f(A)\psi = f(\lambda)\psi.$$

Hint: Use Exercise 5.

7. Suppose $K \subset \mathbb{R}$ is a compact set and μ is a finite measure on K . Let A be the bounded operator on $L^2(K, \mu)$ given by

$$(A\psi)(\lambda) = \lambda\psi(\lambda).$$

Now suppose that B is a bounded operator on $L^2(K, \mu)$ that commutes with A .

- (a) Let $\phi = B\mathbf{1}$, where $\mathbf{1}$ denotes the constant function, so that $\phi \in L^2(K, \mu)$. Show that for all continuous functions ψ on K , we have $B\psi = \phi\psi$.
 - (b) Using Exercise 3, show that for all bounded, Borel-measurable functions ψ on K , we have $B\psi = \phi\psi$.
 - (c) Show that ϕ is essentially bounded (i.e., bounded outside a set of μ -measure zero). Conclude that $B\psi = \phi\psi$ for all $\psi \in L^2(K, \mu)$.
8. If $A \in \mathcal{B}(\mathbf{H})$ is self-adjoint, define $U(t) \in \mathcal{B}(\mathbf{H})$ by $U(t) = \exp\{itA\}$ for each $t \in \mathbb{R}$, where the exponential is defined by the functional calculus for A .
- (a) Show that $U(t)$ is unitary for all t and that $U(s)U(t) = U(s + t)$. (A family of operators with this property is called a *one-parameter unitary group*.)
 - (b) Show that the map $t \mapsto U(t)$ is continuous in the operator norm topology.
 - (c) Give an example of a one-parameter unitary group on a Hilbert space that is not continuous in the operator norm topology.

See Sect. 10.2 for more on one-parameter unitary groups.