10. The Spectral Theorem

The big moment has arrived, and we are now ready to prove several versions of the spectral theorem for normal operators in Hilbert spaces. Throughout this chapter, it should be helpful to compare our results with the more familiar special case when the Hilbert space is finite-dimensional. In this setting, the spectral theorem says that every normal matrix $T \in \mathbb{C}^{n\times n}$ can be diagonalized by a unitary transformation. This can be rephrased as follows: There are numbers $z_j \in \mathbb{C}$ (the eigenvalues) and orthogonal projections $P_j \in B(\mathbb{C}^n)$ such that $T = \sum_{j=1}^m z_j P_j$. The subspaces $R(P_j)$ are orthogonal to each other. From this representation of T, it is then also clear that P_j is the projection onto the eigenspace belonging to z_j.

In fact, we have already proved one version of the (general) spectral theorem: The Gelfand theory of the commutative C^*-algebra $A \subseteq B(H)$ that is generated by a normal operator $T \in B(H)$ provides a *functional calculus:* We can define $f(T)$, for $f \in C(\sigma(T))$ in such a way that the map $C(\sigma(T)) \to A$, $f \mapsto f(T)$ is an isometric $*$-isomorphism between C^*-algebras, and this is the spectral theorem in one of its many disguises! See Theorem 9.13 and the discussion that follows. As a warm-up, let us use this material to give a quick proof of the result about normal matrices $T \in \mathbb{C}^{n\times n}$ that was stated above.

Consider the C^*-algebra $A \subseteq \mathbb{C}^{n\times n}$ that is generated by T. Since T is normal, A is commutative. By Theorem 9.13, $A \cong C(\sigma(T)) = C(\{z_1, \ldots, z_m\})$, where $z_1, \ldots, z_m$ are the eigenvalues of T. We also use the fact that by Theorem 9.16, $\sigma_A(T) = \sigma_{B(H)}(T)$.

All subsets of the discrete space $\{z_1, \ldots, z_m\}$ are open, and thus all functions $f : \{z_1, \ldots, z_m\} \to \mathbb{C}$ are continuous. We will make use of the functional calculus notation: $f(T) \in A$ will denote the operator that corresponds to the function f under the isometric $*$-isomorphism that sends the identity function $\mathrm{id}(z) = z$ to $T \in A$. Write $f_j = \chi_{\{z_j\}}$ and let $P_j = f_j(T)$. Since $\overline{f_j} = f_j$ and $f_j^2 = f_j$, we also have $P_j^* = P_j$ and $P_j^2 = P_j$, so each P_j is an orthogonal projection by Theorem 6.5. Furthermore, $f_j f_k = 0$ if $j \neq k$, so $P_j P_k = 0$, and thus

$$\langle P_j x, P_k y\rangle = \langle x, P_j P_k y\rangle = 0$$

for all $x, y \in H$ if $j \neq k$. This says that $R(P_j) \perp R(P_k)$ for $j \neq k$. Also, $P_1 + \ldots + P_m = 1$ because we have the same identity for the f_j's. It follows that $\bigoplus_{j=1}^m R(P_j) = H = \mathbb{C}^n$. Finally, since $\mathrm{id} = \sum_{j=1}^m z_j f_j$, we obtain the representation $T = \sum_{j=1}^m z_j P_j$, as asserted.

On infinite-dimensional Hilbert spaces, we have a continuous analog of this representation: every normal $T \in B(H)$ can be written as $T =$

$\int z\, dP(z)$. We first need to address the question of how such an integral can be meaningfully defined. We will also switch to the more common symbol E (rather than P) for these "measures" (if that's what they are).

Definition 10.1. Let $\mathcal{M}$ be a σ-algebra on a set Ω, and let H be a Hilbert space. A *resolution of the identity* (or *spectral resolution*) on $(\Omega, \mathcal{M})$ is a map $E : \mathcal{M} \to B(H)$ with the following properties:
(1) Each $E(\omega)$ ($\omega \in \mathcal{M}$) is a projection;
(2) $E(\emptyset) = 0$, $E(\Omega) = 1$;
(3) $E(\omega_1 \cap \omega_2) = E(\omega_1)E(\omega_2) \quad (\omega_1, \omega_2 \in \mathcal{M})$;
(4) For all $x, y \in H$, the set function $\mu_{x,y}(\omega) = \langle x, E(\omega)y\rangle$ is a complex measure on $(\Omega, \mathcal{M})$. If Ω is a locally compact Hausdorff space and $\mathcal{M} = \mathcal{B}$ is the Borel σ-algebra, then we also demand that every $\mu_{x,y}$ is a regular (Borel) measure.

We can think of E as a projection valued measure (of sorts) on $(\Omega, \mathcal{M})$: the "measure" $E(\omega)$ of a set $\omega \in \mathcal{M}$ is a projection. The $E(\omega)$ are also called *spectral projections.*

Let's start out with some quick observations. For every $x \in H$, we have

$$\mu_{x,x}(\omega) = \langle x, E(\omega)x\rangle = \langle x, E(\omega)^2 x\rangle = \langle E(\omega)x, E(\omega)x\rangle = \|E(\omega)x\|^2,$$

so $\mu_{x,x}$ is a finite positive measure with $\mu_{x,x}(\Omega) = \|x\|^2$. Property (3) implies that any two spectral projections $E(\omega_1)$, $E(\omega_2)$ commute. Moreover, if $\omega_1 \subseteq \omega_2$, then $R(E(\omega_1)) \subseteq R(E(\omega_2))$. If $\omega_1 \cap \omega_2 = \emptyset$, then $R(E(\omega_1)) \perp R(E(\omega_1))$, as the following calculation shows:

$$\langle E(\omega_1)x, E(\omega_2)y\rangle = \langle x, E(\omega_1)E(\omega_2)y\rangle = \langle x, E(\omega_1 \cap \omega_2)y\rangle = 0$$

for arbitrary $x, y \in H$.

E is *finitely additive:* If $\omega_1, \ldots, \omega_n \in \mathcal{M}$ are disjoint sets, then $E\left(\bigcup_{j=1}^n \omega_j\right) = \sum_{j=1}^n E(\omega_j)$. To prove this, notice that (4) implies that

$$\begin{aligned}\langle x, E\left(\bigcup_{j=1}^n \omega_j\right) y\rangle &= \mu_{x,y}\left(\bigcup_{j=1}^n \omega_j\right) = \sum_{j=1}^n \mu_{x,y}(\omega_j)\\ &= \sum_{j=1}^n \langle x, E(\omega_j)y\rangle = \langle x, \sum_{j=1}^n E(\omega_j)y\rangle\end{aligned}$$

for arbitrary $x, y \in H$, and this gives the claim.

Is E also σ-additive (as it ought to be, if we are serious about interpreting E as a new sort of measure)? In other words, if $\omega_n \in \mathcal{M}$ are disjoint sets, does it follow that

$$E\left(\bigcup_{n\in\mathbb{N}} \omega_n\right) = \sum_{n=1}^{\infty} E(\omega_n). \tag{10.1}$$

The answer to this question depends on how one defines the right-hand side of (10.1). We observe that if $E(\omega_n) \neq 0$ for infinitely many n, then this series can never be convergent in operator norm. Indeed, $\|E(\omega_n)\| = 1$ if $E(\omega_n) \neq 0$, and thus the partial sums do not form a Cauchy sequence. However, (10.1) will hold if we are satisfied with strong operator convergence: We say that $T_n \in B(H)$ *converges strongly* to $T \in B(H)$ (notation: $T_n \xrightarrow{s} T$) if $T_n x \to Tx$ for all $x \in H$.

To prove that (10.1) holds in this interpretation, fix $x \in H$ and use the fact that the $E(\omega_n)x$ form an orthogonal system (because the ranges of the projections are orthogonal subspaces for disjoint sets). We normalize the non-zero vectors: let $y_n = \frac{E(\omega_n)x}{\|E(\omega_n)x\|}$ if $E(\omega_n)x \neq 0$. Then the y_n form an ONS, and thus, by Theorem 5.15, the series $\sum \langle y_n, x\rangle y_n = \sum E(\omega_n)x$ converges. Now if $y \in H$ is arbitrary, then the continuity of the scalar product and the fact that $\mu_{x,y}$ is a complex measure show that

$$\langle y, \sum_{n=1}^{\infty} E(\omega_n)x\rangle = \sum_{n=1}^{\infty}\langle y, E(\omega_n)x\rangle = \langle y, E\left(\bigcup_{n\in\mathbb{N}} \omega_n\right) x\rangle.$$

Since this holds for every $y \in H$, it follows that $\sum_{n=1}^{\infty} E(\omega_n)x = E\left(\bigcup_{n\in\mathbb{N}} \omega_n\right) x$, and this is (10.1), with the series interpreted as a strong operator limit.

Definition 10.2. A set $N \in \mathcal{M}$ with $E(N) = 0$ is called an *E-null set.* We define $L^\infty(\Omega, E)$ as the set of equivalence classes of measurable, essentially bounded functions $f : \Omega \to \mathbb{C}$. Here, $f \sim g$ if f and g agree off an E-null set. Also, as usual, we say that f is essentially bounded if $|f(x)| \leq M$ $(x \in \Omega \setminus N)$ for some $M \geq 0$ and some E-null set $N \subseteq \Omega$.

Exercise 10.1. Prove that a countable union of E-null sets is an E-null set.

Recall that for an arbitrary positive measure μ on X, the space $L^\infty(X, \mu)$ only depends on what the μ-null sets are and not on the specific choice of the measure μ. For this reason and because of Exercise 10.1, we can also, and in fact without any difficulty, introduce L^∞ spaces that are based on resolutions of the identity. These spaces have the same basic properties: $L^\infty(\Omega, E)$ with the essential supremum of $|f|$ as the norm and the involution $f^*(x) = \overline{f(x)}$ is a commutative

C^*-algebra. The spectrum of a function $f \in L^\infty(\Omega, E)$ is its essential range.

Exercise 10.2. Write down precise definitions of the essential supremum and the essential range of a function $f \in L^\infty(\Omega, E)$.

We would like to define an integral $\int_\Omega f(t)\, dE(t)$ for $f \in L^\infty(\Omega, E)$. This integral should be an operator from $B(H)$, and it also seems reasonable to demand that

$$\langle x, \left(\int_\Omega f(t)\, dE(t) \right) y \rangle = \int_\Omega f(t)\, d\mu_{x,y}(t)$$

for all $x, y \in H$. It is clear that this condition already suffices to uniquely determine $\int_\Omega f(t)\, dE(t)$, should such an operator indeed exist. As for existence, we have the following result; it will actually turn out that the integral with respect to a resolution of the identity has many other desirable properties, too.

Theorem 10.3. *Let E be a resolution of the identity. Then there exists a unique map $\Psi : L^\infty(\Omega, E) \to A$ onto a C^*-subalgebra $A \subseteq B(H)$, such that*

$$\langle x, \Psi(f) y \rangle = \int_\Omega f(t)\, d\mu_{x,y}(t) \tag{10.2}$$

for all $f \in L^\infty(\Omega, E)$, $x, y \in H$. Moreover, Ψ is an isometric $*$-isomorphism from $L^\infty(\Omega, E)$ onto A, and

$$\|\Psi(f)x\|^2 = \int_\Omega |f(t)|^2\, d\mu_{x,x}(t). \tag{10.3}$$

So we can (and will) define $\int_\Omega f(t)\, dE(t) := \Psi(f)$. Let us list the properties of the integral that are guaranteed by Theorem 10.3 one more time, using this new notation:

$$\int (f+g)\, dE = \int f\, dE + \int g\, dE, \quad \int (cf)\, dE = c \int f\, dE,$$
$$\int fg\, dE = \int f\, dE \int g\, dE,$$
$$\left(\int f\, dE \right)^* = \int \overline{f}\, dE, \quad \left\| \int f\, dE \right\| = \|f\|_\infty$$

The multiplicativity of the integral (see the second line) may seem a bit strange at first, but it becomes plausible again if we recall that all $E(\omega)$ are projections.

Exercise 10.3. Show that $\sum_{j=1}^m f_j P_j \sum_{j=1}^m g_j P_j = \sum_{j=1}^m f_j g_j P_j$ if the P_j are projections with orthogonal ranges, as at the beginning of this chapter, and $f_j, g_j \in \mathbb{C}$.

Proof. This is not a particularly short proof, but it follows a standard pattern. First of all, we certainly know how we want to define $\int f\, dE$ for *simple functions* $f \in L^\infty(\Omega, E)$, that is, functions of the form $f = \sum_{j=1}^n c_j \chi_{\omega_j}$ with $c_j \in \mathbb{C}$ and $\omega_j \in \mathcal{M}$. For such an f, put $\Psi(f) = \sum_{j=1}^n c_j E(\omega_j)$. For $x, y \in H$, we then have

$$\langle x, \Psi(f)y\rangle = \sum_{j=1}^n c_j \langle x, E(\omega_j)y\rangle = \sum_{j=1}^n c_j \mu_{x,y}(\omega_j) = \int_\Omega f(t)\, d\mu_{x,y}(t);$$

this is (10.2) for simple functions f, and this identity also confirms that $\Psi(f)$ was indeed well defined ($\Psi(f)$ is determined by the *function* f, and it is independent of the particular representation of f that was chosen to form $\Psi(f)$).

We also have $\Psi(f)^* = \sum \overline{c_j} E(\omega_j) = \Psi(\overline{f})$, and if $g = \sum_{k=1}^m d_k \chi_{\omega'_k}$ is a second simple function, then

$$\Psi(f)\Psi(g) = \sum_{j,k} c_j d_k E(\omega_j) E(\omega'_k) = \sum_{j,k} c_j d_k E(\omega_j \cap \omega'_k) = \Psi(fg).$$

For the last equality, we use the fact that fg is another simple function, with representation $fg = \sum_{j,k} c_j d_k \chi_{\omega_j \cap \omega'_k}$. Similar arguments show that Ψ is linear (on simple functions). Finally, (10.3) (for simple functions) follows from the identity $\Psi(f)^*\Psi(f) = \Psi(\overline{f})\Psi(f) = \Psi(|f|^2)$:

$$\|\Psi(f)x\|^2 = \langle x, \Psi(f)^*\Psi(f)x\rangle = \langle x, \Psi(|f|^2)x\rangle = \int_\Omega |f(t)|^2\, d\mu_{x,x}(t)$$

This also implies that $\|\Psi(f)x\|^2 \le \|f\|_\infty^2 \|x\|^2$, so $\|\Psi(f)\| \le \|f\|$. On the other hand, the sets ω_j in the representation $f = \sum c_j \chi_{\omega_j}$ can be taken to be disjoint (just take $\omega_j = f^{-1}(\{c_j\})$). Now if $E(\omega_j) \ne 0$, then there exists $x \in R(E(\omega_j))$, $x \ne 0$. Clearly, $\Psi(f)x = c_j x$, and since $\|f\|_\infty = \max_{j:E(\omega_j)\ne 0} |c_j|$, we now see that $\|\Psi(f)\| = \|f\|$. So Ψ is isometric (on simple functions).

We now want to extend these results to arbitrary functions $f \in L^\infty(\Omega, E)$ by using an approximation procedure.

Exercise 10.4. Let $f \in L^\infty(\Omega, E)$. Show that there exists a sequence of *simple* functions $f_n \in L^\infty(\Omega, E)$ with $\|f_n - f\| \to 0$.

Let $f \in L^\infty(\Omega, E)$ and pick an approximating sequence f_n of simple functions, as in Exercise 10.4. Notice that $\Psi(f_n)$ converges in $B(H)$:

indeed,

$$\|\Psi(f_m) - \Psi(f_n)\| = \|\Psi(f_m - f_n)\| = \|f_m - f_n\|,$$

so this is a Cauchy sequence. The same argument shows that the limit is independent of the specific choice of the approximating sequence, so we can define $\Psi(f) := \lim \Psi(f_n)$. The continuity of the scalar product gives

$$\langle x, \Psi(f)y\rangle = \lim_{n\to\infty} \langle x, \Psi(f_n)y\rangle = \lim_{n\to\infty} \int_\Omega f_n(t)\, d\mu_{x,y}(t).$$

Every E-null set is a $|\mu_{x,y}|$-null set, so f_n converges $\mu_{x,y}$-almost everywhere to f. Moreover, $|f_n| \le \|f_n\|_\infty \le C$ off an E-null set, so again $\mu_{x,y}$-almost everywhere. The constant function C lies in $L^1(\Omega, d|\mu_{x,y}|)$ because $|\mu_{x,y}|$ is a finite measure. We have just verified the hypotheses of the Dominated Convergence Theorem. It follows that $\lim_{n\to\infty} \int_\Omega f_n\, d\mu_{x,y} = \int_\Omega f\, d\mu$, and we obtain (10.2) (for arbitrary $f \in L^\infty(\Omega, E)$).

Exercise 10.5. Establish (10.3) in a similar way.

The remaining properties follow easily by passing to limits. For example, if $f, g \in L^\infty$, pick approximating simple functions f_n, g_n and use the continuity of the multiplication to deduce that

$$\begin{aligned}\Psi(f)\Psi(g) &= \lim \Psi(f_n) \lim \Psi(g_n) = \lim \Psi(f_n)\Psi(g_n)\\ &= \lim \Psi(f_n g_n) = \Psi(fg).\end{aligned}$$

In the last step, we use the fact that $f_n g_n$ is a sequence of simple functions that converges to fg in the norm of $L^\infty(\Omega, E)$.

Exercise 10.6. Prove at least two more properties of Ψ (Ψ linear, isometric, $\Psi(f)^* = \Psi(\overline{f})$) in this way.

Finally, since Ψ is an isometry, its image $A = \Psi(L^\infty(\Omega, E))$ is closed (compare the proof of Proposition 4.3), and it is also a subalgebra that is closed under the involution $*$ because Ψ is a $*$-homomorphism. □

We now have the tools to prove the next version of the Spectral Theorem (the first version being the existence of a functional calculus for normal operators). We actually obtain a more abstract version for a whole algebra of operators from our machinery; we discuss this first and then specialize to a single operator later on, in Theorem 10.5.

Theorem 10.4. *Suppose $A \subseteq B(H)$ is a commutative C^*-subalgebra of $B(H)$. Let Δ be its maximal ideal space.*

(a) There exists a unique resolution of the identity on the Borel sets of Δ (with its Gelfand topology) such that

$$T = \int_\Delta \widehat{T}(t)\, dE(t) \tag{10.4}$$

for all $T \in A$.

Moreover, E has the following additional properties:
(b) $B = \{\int_\Delta f(t)\, dE(t) : f \in L^\infty(\Omega, E)\}$ is a commutative C^*-algebra satisfying $A \subseteq B \subseteq B(H)$.
(c) The finite linear combinations of the $E(\omega)$, $\omega \in \mathcal{M}$ are dense in B.
(d) If $\omega \subseteq \Delta$ is a non-empty open set, then $E(\omega) \neq 0$.

Proof. By the Gelfand-Naimark Theorem, $A \cong C(\Delta)$. We will now use the Riesz Representation Theorem: $C(\Delta)^* = \mathcal{M}(\Delta)$, the space of regular complex Borel measures on Δ. See Example 4.2. The uniqueness of E follows immediately from this: If E satisfies (10.4), then $\int_\Delta \widehat{T}(t)\, d\mu_{x,y}(t) = \langle x, Ty\rangle$, and every continuous function on Δ is of the form $\widehat{T}$ for some $T \in A$, so the functionals (on $C(\Delta)$) associated with the measures $\mu_{x,y}$ and thus also the measures themselves are already determined by (10.4). Since $x, y \in H$ are arbitrary here, E itself is determined by (10.4).

To prove existence of E, we fix $x, y \in H$ and consider the map $C(\Delta) \to \mathbb{C}$, $\widehat{T} \mapsto \langle x, Ty\rangle$. Since the inverse of the Gelfand transform, $\widehat{T} \mapsto T$, is linear, this map is linear, too, and also bounded, as we see from

$$|\langle x, Ty\rangle| \leq \|x\|\,\|Ty\| \leq \|x\|\,\|T\|\,\|y\| = \|x\|\,\|y\|\,\|\widehat{T}\|_\infty .$$

By the Riesz Representation Theorem, there is a regular complex Borel measure on Δ (call it $\mu_{x,y}$) such that

$$\langle x, Ty\rangle = \int_\Delta \widehat{T}(t)\, d\mu_{x,y}(t) \tag{10.5}$$

for all $T \in A$. Our goal is to construct a resolution of the identity E for which $\langle x, E(\omega)y\rangle = \mu_{x,y}(\omega)$. That will finish the proof of part (a).

As a function of x, y, $\langle x, Ty\rangle$ is sesquilinear. From this, it follows that that $(x, y) \mapsto \mu_{x,y}$ is sesquilinear, too. This means that $\mu_{x+y,z} = \mu_{x,z} + \mu_{y,z}$, $\mu_{cx,y} = \overline{c}\mu_{x,y}$, and $\mu_{x,y}$ is linear in y.

Exercise 10.7. Prove this claim.

If now $f : \Delta \to \mathbb{C}$ is a bounded measurable function, then $(x, y) \mapsto \int_\Delta f(t)\, d\mu_{x,y}(t)$ defines another sesquilinear form. In fact, this form is bounded in the sense that

$$\left| \int_{\Delta} f(t)\, d\mu_{x,y}(t) \right| \le \left(\sup_{t\in\Delta} |f(t)| \right) |\mu_{x,y}|(\Delta) \le \left(\sup_{t\in\Delta} |f(t)| \right) \|x\|\,\|y\|.$$

By Exercise 6.10, there is a unique operator $\Phi(f) \in B(H)$, such that

$$\langle x, \Phi(f)y\rangle = \int_{\Delta} f(t)\, d\mu_{x,y}(t)$$

for all $x, y \in H$. If $f \in C(\Delta)$ here, then a comparison with (10.5) shows that $\Phi(f) = T$, where $T \in A$ is the unique operator with $\widehat{T} = f$. Now

$$\int_{\Delta} \widehat{T}\, d\mu_{x,y} = \langle x, Ty\rangle = \overline{\langle y, T^*x\rangle} = \overline{\int_{\Delta} \widehat{T^*}\, d\mu_{y,x}} = \int_{\Delta} \widehat{T}\, d\overline{\mu_{y,x}},$$

and this holds for all functions $\widehat{T} \in C(\Delta)$, so we conclude that $\mu_{x,y} = \overline{\mu_{y,x}}$, where, as expected, the measure $\overline{\nu}$ is defined by $\overline{\nu}(\omega) = \overline{\nu(\omega)}$. But then we can use this for integrals of arbitrary bounded Borel functions f:

$$\langle x, \Phi(\overline{f})y\rangle = \int_{\Delta} \overline{f}\, d\mu_{x,y} = \overline{\int_{\Delta} f\, d\mu_{y,x}} = \overline{\langle y, \Phi(f)x\rangle} = \langle \Phi(f)x, y\rangle,$$

so $\Phi(f)^* = \Phi(\overline{f})$. Next, for $S, T \in A$, we have

$$\int_{\Delta} \widehat{S}\widehat{T}\, d\mu_{x,y} = \int_{\Delta} (ST)\widehat{\ }d\mu_{x,y} = \langle x, STy\rangle = \int_{\Delta} \widehat{S}\, d\mu_{x,Ty},$$

so $\widehat{T}\, d\mu_{x,y} = d\mu_{x,Ty}$. Again, we can apply this to integrals of arbitrary bounded Borel functions f: $\int f\widehat{T}\, d\mu_{x,y} = \int f\, d\mu_{x,Ty}$, and this implies that

$$\int_{\Delta} f\widehat{T}\, d\mu_{x,y} = \langle x, \Phi(f)Ty\rangle = \langle \Phi(f)^*x, Ty\rangle = \int_{\Delta} \widehat{T}\, d\mu_{\Phi(f)^*x,y}.$$

Since $\widehat{T} \in C(\Delta)$ is arbitrary here, this says that $f\, d\mu_{x,y} = d\mu_{\Phi(f)^*x,y}$, so $\int fg\, d\mu_{x,y} = \int g\, d\mu_{\Phi(f)^*x,y}$ for all bounded Borel functions g. Now $\int fg\, d\mu_{x,y} = \langle x, \Phi(fg)y\rangle$ and

$$\int_{\Delta} g\, d\mu_{\Phi(f)^*x,y} = \langle \Phi(f)^*x, \Phi(g)y\rangle = \langle x, \Phi(f)\Phi(g)y\rangle,$$

so we finally obtain the desired conclusion that $\Phi(fg) = \Phi(f)\Phi(g)$.

We can now define $E(\omega) = \Phi(\chi_\omega)$. I claim that E is a resolution of the identity. Clearly, by construction,

$$\langle x, E(\omega)y\rangle = \langle x, \Phi(\chi_\omega)y\rangle = \int_{\Delta} \chi_\omega\, d\mu_{x,y} = \mu_{x,y}(\omega),$$

as required. This also verifies (4) from Definition 10.1. It remains to check the conditions (1)–(3).

Notice that $E(\omega)^* = \Phi(\chi_\omega)^* = \Phi(\overline{\chi_\omega}) = \Phi(\chi_\omega) = E(\omega)$, so $E(\omega)$ is self-adjoint. Similarly, $E(\omega)^2 = \Phi(\chi_\omega)^2 = \Phi(\chi_\omega^2) = \Phi(\chi_\omega) = E(\omega)$. By Theorem 6.5, $E(\omega)$ is a projection, so (1) holds. A similar computation lets us verify (3). Finally, moving on to (2), it is clear that $E(\emptyset) = \Phi(0) = 0$, and $E(\Delta) = \Phi(1)$. Now the constant function 1 is continuous, so, as observed above, $\Phi(1)$ is the operator whose Gelfand transform is identically equal to one, but this is the identity operator $1 \in A \subseteq B(H)$ (the multiplicative unit of A and $B(H)$). So $E(1) = 1$, as desired.

(b) We know from Theorem 10.3 that B is a C^*-subalgebra of $B(H)$, and since continuous functions are in $L^\infty(\Delta, E)$, we clearly have $B \supseteq A$.

(c) This is immediate from the way the integral $\int f\, dE$ was constructed, in the proof of Theorem 10.3.

(d) Let $\omega \subseteq \Delta$ be a non-empty open set. Pick $t_0 \in \omega$ and use Urysohn's Lemma to find a continuous function f with $f(t_0) = 1$, $f = 0$ on ω^c. Then $f = \widehat{T}$ for some $T \in A$, and if we had $E(\omega) = 0$, then $T = \int_\Delta \widehat{T}\, dE = 0$, but this is impossible because $\widehat{T} = f$ is not the zero function. □

We now specialize to (algebras generated by) a single normal operator.

Theorem 10.5 (The Spectral Theorem for normal operators)**.** *Let $T \in B(H)$ be a normal operator. Then there exists a unique resolution of the identity E on the Borel sets of $\sigma(T)$ such that*

$$T = \int_{\sigma(T)} z\, dE(z). \tag{10.6}$$

Proof. Consider, as usual, the commutative C^*-algebra $A \subseteq B(H)$ that is generated by T. Existence of E now follows from Theorem 10.4(a) because we can make the following identifications: By Theorems 9.12 and 9.13, Δ_A is homeomorphic to $\sigma(T)$, and $A \cong C(\sigma(T))$. Here we may interpret $\sigma(T)$ as $\sigma_{B(H)}(T)$ because $\sigma_A(T)$ is the same set by Theorem 9.16. From a formal point of view, perhaps the most satisfactory argument runs as follows: Reexamine the proof of Theorem 10.4 to confirm that we obtain the representation $T = \int_K f\, dE$ as soon as we have an isometric $*$-isomorphism between A and $C(K)$ that sends T to f (it is not essential that this isomorphism is specifically the Gelfand transform). In the case at hand, $A \cong C(\sigma(T))$, by Theorem 9.13, and the corresponding isomorphism sends T to $\mathrm{id}(z) = z$, so we obtain (10.6). For later use, we also record that, by the same argument,

$f(T) = \int_{\sigma(T)} f(z)\, dE(z)$ for all $f \in C(\sigma(T))$, where $f(T) \in A$ is defined as in Chapter 9; see especially the discussion following Theorem 9.13.

Let us now prove uniqueness of E. By Theorem 10.3, if (10.6) holds, then also $p(T,T^*) = \int_{\sigma(T)} p(z,\overline{z})\, dE(z)$ for all polynomials p in two variables. When viewed as functions of z only, this set

$$\{f : \sigma(T) \to \mathbb{C} : f(z) = p(z,\overline{z}),\ p \text{ polynomial in two variables}\}$$

satisfies the hypotheses of the Stone-Weierstraß Theorem. Therefore, if $f \in C(\sigma(T))$ is arbitrary, there are polynomials p_n such that $\|f(z) - p_n(z,\overline{z})\|_\infty \to 0$. Alternatively, this conclusion can also be obtained from the fact that T generates A, so $\{p(T,T^*)\}$ is dense in A, and we can then move things over to $C(\sigma(T))$.

The Dominated Convergence Theorem shows that

$$\int_{\sigma(T)} f(z)\, d\mu_{x,y}(z) = \lim_{n\to\infty} \int_{\sigma(T)} p_n(z,\overline{z})\, d\mu_{x,y}(z) = \lim_{n\to\infty} \langle x, p_n(T,T^*)y\rangle .$$

So the measures $\mu_{x,y}$ and thus also E itself are uniquely determined. □

This proof has also established the following fact, which we state again because it will prove useful in the sequel:

Proposition 10.6. *If E is the spectral resolution of a normal operator $T \in B(H)$, as in the Spectral Theorem, then $E(U \cap \sigma(T)) \neq 0$ for all open sets $U \subseteq \mathbb{C}$ with $U \cap \sigma(T) \neq \emptyset$.*

This follows from Theorem 10.4(d) and our identification of Δ_A with $\sigma(T)$.

We introduce some new notation. It will occasionally be convenient to write $d\langle x, E(z)y\rangle$ for the measure $d\mu_{x,y}(z)$. Similarly, $d\langle x, E(z)x\rangle$ and $d\|E(z)x\|^2$ both refer to the measure $d\mu_{x,x}(z)$. This notation is reasonable because $\langle x, E(\omega)x\rangle = \|E(\omega)x\|^2$.

We can now also extend the *functional calculus* from Chapter 9. More precisely, for a normal $T \in B(H)$ and $f \in L^\infty(\sigma(T), E)$, where E is the resolution of the identity of T, as in the Spectral Theorem, let

$$f(T) := \int_{\sigma(T)} f(z)\, dE(z). \tag{10.7}$$

As observed above, in the proof of Theorem 10.5, this is consistent with our earlier definition of $f(T)$ for $f \in C(\sigma(T))$ from Chapter 9.

By Theorem 10.3, the functional calculus $f \mapsto f(T)$ is an isometric $*$-isomorphism between $L^\infty(\sigma(T), E)$ and a subalgebra of $B(H)$. Note also that if $p(z)$ is a polynomial, $p(z) = \sum_{j=0}^n c_j z^j$, then $p(T)$ could have been defined directly as $p(T) = \sum_{j=0}^n c_j T^j$, and the functional calculus gives the same result. A similar remark applies to functions of the form $p(z, \overline{z})$.

We state the basic properties of the functional calculus one more time:

$$(cf + dg)(T) = cf(T) + dg(T), \quad (fg)(T) = f(T)g(T) = g(T)f(T)$$

$$f(T)^* = \overline{f}(T), \quad \|f(T)\| = \|f\|_\infty, \quad \|f(T)x\|^2 = \int_{\sigma(T)} |f(z)|^2 \, d\|E(z)x\|^2$$

Moreover, if f is continuous, then we have the *spectral mapping theorem:* $\sigma(f(T)) = f(\sigma(T))$. This was discussed in Exercise 9.16.

We want to prove still another version of the Spectral Theorem. This last version will be an analog of the statement: a normal matrix can be diagonalized by a unitary transformation. We will needs sums of Hilbert spaces to formulate this result, so we discuss this topic first. If $H_1, \ldots, H_n$ are Hilbert spaces, then we can construct a new Hilbert space $H = \bigoplus_{j=1}^n H_j$, as follows: As a vector space, H is the sum of the vector spaces H_j, and if $x, y \in H$, say $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, then we define $\langle x, y\rangle = \sum_{j=1}^n \langle x_j, y_j\rangle_{H_j}$.

Exercise 10.8. Verify that this defines a scalar product on H and that H is complete with respect to the corresponding norm.

Note that each H_j can be naturally identified with a closed subspace of H, by sending $x_j \in H_j$ to $x = (0, \ldots, 0, x_j, 0, \ldots 0)$. In fact, the H_j, viewed in this way as subspaces of H, are pairwise orthogonal. Conversely, if H is a Hilbert space and the H_j are orthogonal subspaces of H, then $\bigoplus H_j$ can be naturally identified with a subspace of H (by mapping (x_j) to $\sum x_j$).

An analogous construction works for infinitely many summands H_α, $\alpha \in I$. We now define H to be the set of vectors $x = (x_\alpha)_{\alpha\in I}$ ($x_\alpha \in H_\alpha$) that satisfy $\sum_{\alpha\in I} \|x_\alpha\|^2 < \infty$. If I is uncountable, then, as usual, this means that $x_\alpha \neq 0$ for only countably many α and the corresponding series is required to converge. We can again define $\langle x, y\rangle = \sum_{\alpha\in I}\langle x_\alpha, y_\alpha\rangle$; the convergence of this series follows from the definition on H and the Cauchy-Schwarz inequality for both the individual scalar products and then also the sum over $\alpha \in I$.

Exercise 10.9. Again, prove that this defines a scalar product and that H is a Hilbert space.

Theorem 10.7 (Spectral representation of normal operators)**.** *Let $T \in B(H)$ be a normal operator. Then there exist a collection $\{\rho_\alpha : \alpha \in I\}$ of finite positive Borel measures on $\sigma(T)$ and a unitary map $U : H \to$*

$\bigoplus_{\alpha\in I} L^2(\sigma(T), d\rho_\alpha)$ such that

$$UTU^{-1} = M_z, \qquad (M_z f)_\alpha(z) = z f_\alpha(z).$$

The minimal cardinality of such a set I is called the *spectral multiplicity* of T; if H is separable (as almost all Hilbert spaces that occur in practice are), then I can always taken to be a countable set (say $I = \mathbb{N}$). Sometimes, a finite I will suffice or even an I consisting of just one element, so that T would then be unitarily equivalent to a multiplication operator by the variable on a single $L^2(\rho)$ space.

Exercise 10.10. Let $T \in \mathbb{C}^{n\times n}$ be a normal matrix, with eigenvalues $\sigma(T) = \{z_1, \ldots, z_m\}$. Prove the existence of a spectral representation directly, by providing the details in the following sketch: Choose the ρ_α as counting measures on (subsets of) $\sigma(T)$, and to define U, send a vector $x \in \mathbb{C}^n$ to its expansion coefficients with respect to an ONB consisting of eigenvectors of T.

Exercise 10.11. Use the discussion of the previous Exercise to show that for a normal $T \in \mathbb{C}^{n\times n}$, the spectral multiplicity (as defined above) is the maximal degeneracy of an eigenvalue, or, put differently, it is equal to $\max_{z\in\sigma(T)} \dim N(T - z)$.

The measures ρ_α from Theorem 10.7 are called *spectral measures.* They are *not* uniquely determined by the operator T; Exercise 10.17 below will shed some additional light on this issue.

Proof. For $x \in H$, $x \neq 0$, let

$$H_x = \overline{\{f(T)x : f \in C(\sigma(T))\}}.$$

We also define an operator $U_x : H_x \to L^2(\sigma(T), d\mu_{x,x})$, as follows: For $f \in C(\sigma(T))$, put $U_x^{(0)} f(T)x = f$. Then

$$\|U_x^{(0)} f(T)x\|^2 = \int_{\sigma(T)} |f(z)|^2 \, d\mu_{x,x}(z) = \|f(T)x\|^2,$$

by Theorem 10.3. By Exercise 2.26, the operator $U_x^{(0)} : \{f(T)x\} \to L^2(\mu_{x,x})$ has a unique continuous extension to H_x (call it U_x). Since the norm is continuous, U_x will also be isometric. In particular, $R(U_x)$ is closed, but clearly $R(U_x)$ also contains every continuous function on $\sigma(T)$, and these are dense in $L^2(\sigma(T), d\mu_{x,x})$, so $R(U_x) = L^2(\sigma(T), d\mu_{x,x})$. Summing up: U_x is a unitary map (a linear bijective isometry) from H_x onto $L^2(\sigma(T), d\mu_{x,x})$.

Now let $f \in C(\sigma(T))$ and write $zf(z) = g(z)$. Then

$$U_x T U_x^{-1} f = U_x T f(T)x = U_x g(T)x = g = M_z f,$$

where M_z denotes the operator of multiplication by z (here: in $L^2(\sigma(T), d\mu_{x,x})$). Since these functions f are dense in $L^2(\sigma(T), d\mu_{x,x})$ and both operators $U_xTU_x^{-1}$ and M_z are continuous, it follows that $U_xTU_x^{-1} = M_z$.

We now consider those collections of such spaces $\{H_x : x \in I\}$ for which the individual spaces are orthogonal: $H_x \perp H_y$ if $x, y \in I$, $x \neq y$. One can use Zorn's Lemma to show that there is such collection of H_x spaces with $\bigoplus_{x\in I} H_x = H$. As always, we don't want to discuss the details of this argument. The crucial fact is this: If $\bigoplus_{x\in I} H_x \neq H$, then there is another space H_y that is orthogonal to all H_x ($x \in I$). This can be proved as follows: Just pick an arbitrary $y \in \left(\bigoplus_{x\in I} H_x\right)^{\perp}$, $y \neq 0$. Then $\langle y, g(T)x\rangle = 0$ for all $x \in I$ and continuous functions g. But then it also follows that for all continuous f

$$\langle f(T)y, g(T)x\rangle = \langle y, \overline{f}(T)g(T)x\rangle = 0,$$

because $\overline{f}g$ is another continuous function. So $f(T)y \perp H_x$ and thus $H_y \perp H_x$ by the continuity of the scalar product.

We can now define the unitary map U as $U = \bigoplus_{x\in I} U_x$, where I is chosen such that $\bigoplus_{x\in I} H_x = H$, as discussed in the preceding paragraph. More precisely, by this we mean the following:

$$U : H \to \bigoplus_{x\in I} L^2(\sigma(T), d\mu_{x,x}),$$

and if $y = \sum_{x\in I} y_x$ is the unique decomposition of $y \in H$ into components $y_x \in H_x$, then we put $(Uy)_x = U_x y_x$. This map has the desired properties.

Exercise 10.12. Check this in greater detail.

□

We have now discussed three versions of the Spectral Theorem. We originally obtained the functional calculus for normal operators from the theory of C^*-algebras, especially the Gelfand-Naimark Theorem. This was then used to derive the existence of a spectral resolution E and a spectral representation. Conversely, spectral resolutions can be used to construct (in fact: an extended version of) the functional calculus, and it is also easy to recover E, starting from a spectral representation $UTU^{-1} = M_z$ (we sometimes write this as $T \cong M_z$). We summarize symbolically:

$$\boxed{\text{functional calculus} \quad \Longleftrightarrow \quad T = \int_{\sigma(T)} z\, dE(z) \quad \Longleftrightarrow \quad T \cong M_z}$$

Every version has its merits, and it's good to have all three statements available. Note, however, that the original functional calculus (obtained from the theory of C^*-algebras) becomes superfluous now because we obtain more powerful versions from the other statements (this was already pointed out above).

The spectrum of T will not always be known, and so it is sometimes more convenient to have statements that do not explicitly involve $\sigma(T)$. This is very easy to do: Given E, we can also get a spectral resolution on the Borel sets of $\mathbb{C}$ by simply declaring $E(\mathbb{C} \setminus \sigma(T)) = 0$. Similarly, in a spectral representation, we can think of the ρ_α as measures on $\mathbb{C}$ (with $\rho_\alpha(\mathbb{C} \setminus \sigma(T)) = 0$).

In this case, we can recover the spectrum from the measures ρ_α. We discuss the case of one space $L^2(\mathbb{C}, d\rho)$ and leave the discussion of the effect of the orthogonal sum to an exercise. Given a Borel measure ρ on $\mathbb{C}$, we define its *topological support* as the smallest closed set A that supports ρ in the sense that $\rho(A^c) = 0$. We denote it by $A = \text{top supp } \rho$.

Exercise 10.13. Prove that such a set exists. *Suggestion:* It is tempting to try to define $(\text{top supp } \rho)^c = \bigcup U$, where the union is over all open sets $U \subseteq \mathbb{C}$ with $\rho(U) = 0$. This works, but note that the union will be uncountable, which is a nuisance from a technical point of view because we want to show that it has ρ measure zero.

Proposition 10.8. *If $T = M_z$ on $L^2(\mathbb{C}, d\rho)$, then $\sigma(T) = \text{top supp } \rho$.*

Proof. Abbreviate $S = \text{top supp } \rho$. We must show that $M_z - w$ is invertible in $B(L^2)$ precisely if $w \notin S$. Now if $w \notin S$, then $|w - z| \geq \epsilon > 0$ for ρ-almost every $z \in \mathbb{C}$ (by definition of S), and this implies that $M_{(z-w)^{-1}}$ is a bounded linear operator. Obviously, it is the inverse of $M_z - w$.

Conversely, if $w \in S$, then $\rho(B_n) > 0$ for all $n \in \mathbb{N}$, where $B_n = \{z \in \mathbb{C} : |z - w| < 1/n\}$. Again, this follows from the definition of S. This means that $\|\chi_{B_n}\| > 0$ in $L^2(\mathbb{C}, d\rho)$. Let $f_n = \chi_{B_n}/\|\chi_{B_n}\|$, so $\|f_n\| = 1$. Then $\|(M_z - w)f_n\| < 1/n$, and this shows that $(M_z - w)$ is not invertible: if it were, then it would follow that

$$1 = \|f_n\| = \|(M_z - w)^{-1}(M_z - w)f_n\| \leq C\|(M_z - w)f_n\| < \frac{C}{n},$$

which is absurd. $\square$

As for the orthogonal sum, we have the following result:

Proposition 10.9. *Let H_α be Hilbert spaces, and let $T_\alpha \in B(H_\alpha)$ be normal operators, with $\sup_{\alpha \in I} \|T_\alpha\| < \infty$. Write $H = \bigoplus_{\alpha \in I} H_\alpha$ and* *define* $T : H \to H$ *as follows:* $(Tx)_\alpha = T_\alpha x_\alpha$ *(if* $x = (x_\alpha)_{\alpha\in I}$*). Then* $T \in B(H)$ *and*

$$\sigma(T) = \overline{\bigcup_{\alpha\in I} \sigma(T_\alpha)}.$$

It is customary to write this operator as $T = \bigoplus_{\alpha\in I} T_\alpha$, and actually we already briefly mentioned this notation in the proof of Theorem 10.7. If I is finite, then no closure is necessary in the statement of Proposition 10.9.

The situation of Theorem 10.7 is as discussed in the Proposition, with $T_\alpha = M_z$ for all α. So we can now say that the spectrum of M_z on $\bigoplus L^2(\mathbb{C}, d\rho_\alpha)$ is the closure of the union of the topological supports of the ρ_α.

Exercise 10.14. Prove Proposition 10.9.

The following basic facts are very useful when dealing with spectral representations. They provide further insight into the functional calculus and also a very convenient way of performing these operations once a spectral representation has been found.

Proposition 10.10. *Let* $f : \mathbb{C} \to \mathbb{C}$ *be a bounded Borel function. Then:*
(a) $f(M_z) = M_{f(z)}$*;*
(b) Let $U : H_1 \to H_2$ *be a unitary map and let* $T \in B(H_1)$ *be a normal operator. Then*

$$f(UTU^{-1}) = Uf(T)U^{-1}.$$

Sketch of proof. We argue as in the second part of the proof of Theorem 10.5. First of all, the assertions hold for functions of the type $f(z) = p(z, \overline{z})$, with a polynomial p, because for such functions we have an alternative direct description of $f(T)$, which lets us verify (a), (b) directly. Again, by the Stone-Weierstraß Theorem, these functions are dense in $C(K)$ for compact subsets $K \subseteq \mathbb{C}$. Since $f_n(T) \to f(T)$ in $B(H)$ if $\|f_n - f\|_\infty \to 0$, this gives the claim for continuous functions. Now $\|(f(T) - g(T))x\|^2 = \int |f - g|^2 \, d\mu_{x,x}$ and continuous functions are dense in L^2 spaces. From this, we obtain the statements for arbitrary bounded Borel functions. □

Exercise 10.15. Give a detailed proof by filling in the details.

If T is of the form M_z on $L^2(\mathbb{C}, d\rho)$, as in a spectral representation (where we assume, for simplicity, that there is just one L^2 space), what is the spectral resolution E of this operator? In general, we can recover E from T as $E(A) = \chi_A(T)$, so Proposition 10.10 shows that $E(A) = M_{\chi_A}$ if $A \subseteq \mathbb{C}$ is a Borel set.

Exercise 10.16. Verify directly that this defines a resolution of the identity on the Borel sets of $\mathbb{C}$ (and the Hilbert space $L^2(\mathbb{C}, d\rho)$).

We observed earlier that the spectral measures ρ_α are (in fact: highly) non-unique. The following Exercise helps to clarify the situation. We call two operators $T_j \in B(H_j)$ *unitarily equivalent* if $T_2 = UT_1U^{-1}$ for some unitary map $U : H_1 \to H_2$. So, if we use this terminology, then Theorem 10.7 says that every normal operator is unitarily equivalent to the operator of multiplication by the variable in a sum of spaces $L^2(\mathbb{C}, \rho_\alpha)$.

Exercise 10.17. Consider the multiplication operators $T_1 = M_z^{(\mu)}$ and $T_2 = M_z^{(\nu)}$ on $L^2(\mu)$ and $L^2(\nu)$, respectively, where μ, ν are finite Borel measures on $\mathbb{C}$. Show that T_1, T_2 are unitarily equivalent if and only if μ and ν are equivalent measures (that is, they have the same null sets). *Suggestion:* For one direction, use the fact that μ and ν are equivalent if and only if $d\mu = f\, d\nu$, with $f \in L^1(\nu)$ and $f > 0$ almost everywhere with respect to μ (or ν).

Example 10.1. Let us now discuss the operator $(Tx)_n = x_{n+1}$ on $\ell^2(\mathbb{Z})$. By Exercise 6.7(a), T is unitary, so the Spectral Theorem applies. It is easiest to start out with a spectral representation because this can be guessed directly. Consider the operator

$$F : L^2(S, dx/(2\pi)) \to \ell^2(\mathbb{Z}), \quad (Ff)_n = \frac{1}{2\pi}\int_0^{2\pi} f(e^{ix})e^{inx}\, dx$$

(F as in *Fourier transform*). Here, $S = \{z \in \mathbb{C} : |z| = 1\}$ denotes again the unit circle; when convenient, we also use $x \in [0, 2\pi)$ to parametrize S by writing $z = e^{ix}$. Note that $(Ff)_n = \langle e_n, f\rangle$, with $e_n(z) = z^{-n}$. Since these functions form an ONB (compare Exercise 5.15), Theorem 5.14 shows that F is unitary.

Observe that the function $g(z) = zf(z)$ has Fourier coefficients $(Fg)_n = (Ff)_{n+1}$. In other words, $F^{-1}TF = M_z$, and this is a spectral representation, with $U = F^{-1}$. The spectral measure $dx/(2\pi)$ has the unit circle as its topological support, so $\sigma(T) = S$. Since only one L^2 space is necessary here, the operator T has spectral multiplicity one.

What is the spectral resolution of T? We already identified this spectral resolution on $L^2(S, dx/(2\pi))$, the space from the spectral representation, and we can now map things back to the original Hilbert space $\ell^2(\mathbb{Z})$ by using Proposition 10.10. More specifically,

$$E(A) = \chi_A(T) = \chi_A(FM_zF^{-1}) = F\chi_A(M_z)F^{-1} = FM_{\chi_A(z)}F^{-1}.$$

We can rewrite this if we recall that $(F^{-1}y)(z) = \sum y_n z^{-n}$, so $(M_{\chi_A} F^{-1}y)(z) = \sum y_n \chi_A(z) z^{-n}$ (both series converge in $L^2(S)$), and thus

$$(E(A)y)_n = \sum_{m=-\infty}^{\infty} \widehat{\chi_A}(m-n) y_m,$$

where $\widehat{\chi_A}(k) = 1/(2\pi) \int_0^{2\pi} \chi_A(e^{ix}) e^{ikx}\, dx$. Formally, this follows immediately from the preceding formulae, and for a rigorous argument, we use the fact that $(Ff)_n$ may be interpreted as a scalar product, the continuity of the scalar product and the L^2 convergence of the series that are involved here.

We now prove some general statements that illustrate how the Spectral Theorem helps to analyze normal operators.

Theorem 10.11. *Let $T \in B(H)$ be normal. Then:*
(a) T is self-adjoint $\iff \sigma(T) \subseteq \mathbb{R}$;
(b) T is unitary $\iff \sigma(T) \subseteq S = \{z \in \mathbb{C} : |z| = 1\}$.

The assumption that T is normal is needed here: if, for example, $T = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \in B(\mathbb{C}^2)$, then $\sigma(T) = \{0\} \subseteq \mathbb{R}$, but T is not self-adjoint.

Proof. (a) $\Longrightarrow$: This was established earlier, in Theorem 9.15(a).
$\Longleftarrow$: By the Spectral Theorem and functional calculus,

$$T^* = \int_{\sigma(T)} \overline{z}\, dE(z) = \int_{\sigma(T)} z\, dE(z) = T.$$

(b) $\Longleftarrow$: This follows as in (a) from

$$TT^* = T^*T = \int_{\sigma(T)} z\overline{z}\, dE(z) = \int_{\sigma(T)} dE(z) = 1.$$

$\Longrightarrow$: If $z \in \sigma(T)$, then $E(B_{1/n}(z)) \neq 0$ for all $n \in \mathbb{N}$ by Proposition 10.6, so we can pick $x_n \in R(E(B_{1/n}(z)))$, $\|x_n\| = 1$. Then

$$\begin{aligned}\mu_{x_n,x_n}((B_{1/n}(z))^c) &= \langle x_n, E((B_{1/n}(z))^c) x_n\rangle \\ &= \langle x_n, E((B_{1/n}(z))^c) E(B_{1/n}(z)) x_n\rangle = 0,\end{aligned}$$

so it follows that
(10.8)
$$\big|\|Tx_n\| - |z|\,\|x_n\|\big|^2 \le \|(T-z)x_n\|^2 = \int_{\sigma(T)} |t-z|^2\, d\mu_{x_n,x_n}(t) \le \frac{1}{n^2}.$$

Since $\|Ty\| = \|y\|$ for all $y \in H$ for a unitary operator, this shows that $|z| = 1$, as claimed. $\square$

Theorem 10.12. *If $T \in B(H)$ is normal, then*
$$\|T\| = \sup_{\|x\|=1} |\langle x, Tx\rangle| .$$

Proof. Clearly, $|\langle x, Tx\rangle| \le \|T\|\,\|x\|^2$, so the sup is $\le \|T\|$. On the other hand, we know from Theorem 9.15(b) that $\|T\| = r(T)$, so there exists a $z \in \sigma(T)$ with $|z| = \|T\|$. As in the previous proof, if $\epsilon > 0$ is given, then $E(B_\epsilon(z)) \ne 0$, so we can find an $x \in R(E(B_\epsilon(z)))$, $\|x\| = 1$. Then
$$|\langle x, Tx\rangle - z| = |\langle x, (T-z)x\rangle| = \left| \int (t-z)\, d\|E(t)x\|^2 \right| < \epsilon$$
because (again, as in the previous proof) $\mu_{x,x}((B_\epsilon(z))^c) = 0$ (and $\mu_{x,x}(\mathbb{C}) = \|x\|^2 = 1$). So $\sup|\langle x, Tx\rangle| \ge \|T\| - \epsilon$, and $\epsilon > 0$ is arbitrary here. $\square$

Theorem 10.13. *Let $T \in B(H)$. Then $T \ge 0$ (in the C^*-algebra $B(H)$; see Definition 9.14) if and only if $\langle x, Tx\rangle \ge 0$ for all $x \in H$.*

Proof. If $T \ge 0$, then T is self-adjoint and $\sigma(T) \subseteq [0,\infty)$, so the Spectral Theorem shows that
$$\langle x, Tx\rangle = \int_{[0,\infty)} t\, d\mu_{x,x}(t) \ge 0$$
for all $x \in H$.

Conversely, if this condition holds, then in particular $\langle x, Tx\rangle \in \mathbb{R}$ for all $x \in H$, so $\langle x, T^*x\rangle = \langle Tx, x\rangle = \overline{\langle x, Tx\rangle} = \langle x, Tx\rangle$. Polarization now shows that $\langle x, T^*y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$, that is, $T = T^*$ and T is self-adjoint.

Now if $t > 0$, then
$$t\|x\|^2 = \langle x, tx\rangle \le \langle x, (T+t)x\rangle \le \|x\|\,\|(T+t)x\|,$$
so it follows that
$$\|(T+t)x\| \ge t\|x\|. \tag{10.9}$$
This shows, first of all, that $N(T+t) = \{0\}$. Moreover, we also see from (10.9) that $R(T+t)$ is closed: if $y_n \in R(T+t)$, say $y_n = (T+t)x_n$ and $y_n \to y \in H$, then (10.9) shows that x_n is a Cauchy sequence, so $x_n \to x$ for some $x \in H$ and thus $y = (T+t)x \in R(T+t)$ also, by the continuity of $T+t$. Finally, we observe that $R(T+t)^\perp = N((T+t)^*) = N(T+t) = \{0\}$ (by Theorem 6.2). Putting things together, we see that $R(T+t) = H$, so $T+t$ is bijective and thus $-t \notin \sigma(T)$. This holds for every $t > 0$, so, since T is self-adjoint, $\sigma(T) \subseteq [0,\infty)$ and $T \ge 0$. $\square$

Theorem 10.14. *Let $T \in B(H)$, $T \ge 0$. Then there exists a unique $S \in B(H)$, $S \ge 0$ with $S^2 = T$.*

Proof. Existence is very easy: By the Spectral Theorem, $T = \int_{[0,\infty)} t\, dE(t)$. The operator $S = \int_{[0,\infty)} t^{1/2}\, dE(t)$ has the desired properties (here, $t^{1/2}$ of course denotes the positive square root).

Uniqueness isn't hard either, but more technical, and we just sketch this part: If S_0 is another operator with $S_0 \geq 0$, $S_0^2 = T$, write $S_0 = \int_{[0,\infty)} s\, dE_0(s)$, so $T = \int_{[0,\infty)} s^2\, dE_0(s)$. Now we can run a "substitution" $s^2 = t$ (of sorts) and rewrite this as $T = \int_{[0,\infty)} t\, d\widetilde{E}_0(t)$, where $\widetilde{E}_0(M) = E_0(\{s^2 : s \in M\})$ (this part would need a more serious discussion if a full proof is desired). By the uniqueness of the spectral resolution E (see Theorem 10.5), $\widetilde{E}_0 = E$, and this will imply that $S_0 = S$. □

Exercise 10.18. Let $T \in \mathbb{C}^{n\times n}$ be a normal matrix with n distinct, non-zero eigenvalues. Show that there are precisely 2^n normal (!) matrices $S \in \mathbb{C}^{n\times n}$ with $S^2 = T$.

Exercise 10.19. Recall that $\sigma_p(T)$ was defined as the set of eigenvalues of T; equivalently, $z \in \sigma_p(T)$ precisely if $N(T-z) \neq \{0\}$. Show that if $T \in B(H)$ is normal, then $z \in \sigma_p(T)$ if and only if $E(\{z\}) \neq 0$ (here, as usual, E denotes the spectral resolution of T).

Exercise 10.20. Let $T \in B(H)$ be normal. Show that $z \in \sigma(T)$ if and only if there exists a sequence $x_n \in H$, $\|x_n\| = 1$, such that $(T-z)x_n \to 0$.

Exercise 10.21. Suppose that $T \in B(H)$ is both unitary and self-adjoint. Show that T is of the form $T = 2P - 1$, for some orthogonal projection P. Show also that, conversely, every such operator T is unitary and self-adjoint.
Suggestion: Use the Spectral Theorem and Theorem 10.11 for the first part.

Exercise 10.22. Let $T \in B(H)$. Recall that a closed subspace $M \subseteq H$ is called *invariant* if $TM \subseteq M$, that is, if $Tx \in M$ for all $x \in M$. Call M a *reducing subspace* if both M and $M^\perp$ are invariant. Show that if T is normal with spectral resolution E, then $R(E(B))$ is a reducing subspace for every Borel set $B \subseteq \mathbb{C}$.
Hint: $E(B) = \chi_B(T)$; now use the functional calculus.