The Spectral Theorem.

Alonso Delfín
University of Oregon.

June 7, 2018

Abstract.

In 1963, Paul Halmos wrote the following in a paper:

"*Most students of mathematics learn quite early and most mathematicians remember till quite late that every Hermitian matrix may be put into diagonal form. The spectral theorem is widely and correctly regarded as the generalization of this assertion to operators on Hilbert space. It is unfortunate therefore that even the bare statement of the spectral theorem is widely regarded as somewhat mysterious and deep, and probably inaccessible to the nonspecialist.* [...] *Another reason the spectral theorem is thought to be hard is that its proof is hard*"

The purpose of my talk is the same Halmos' paper has, to try to dispel some of the mystery behind the spectral theorem. I'll go over at least two examples, which should be accessible to everyone. Then, I'll give a rough sketch of the spectral theorem's proof, in which I'll assume some basic knowledge in measure theory and C^*-algebras.

Motivation from Linear Algebra

Recall a matrix $A \in M_n(\mathbb{C})$ is said to be Hermitian (or self adjoint) if it is equal to its conjugate transpose, that is $A^* = A$. We say that a matrix $A \in M_n(\mathbb{C})$ is normal if $A^*A = AA^*$. A matrix $A \in M_n(\mathbb{C})$ is unitary if $A^*A = I_n = AA^*$.

The linear algebraic statement of the spectral theorem says that if $A \in$

$M_n(\mathbb{C})$ is normal, then there is a unitary matrix $U \in M_n(\mathbb{C})$ such that

$$A = U\Lambda U^*$$

where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ for $\{\lambda_1, \ldots, \lambda_n\} = \sigma(A)$.

We are now interested on giving a statement for the spectral theorem for normal operators on $\mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is a separable Hilbert space not necessarily finite dimensional.

Example. Suppose $\mathcal{H}$ is an infinite dimensional, separable Hilbert space. Let $T \in \mathbb{B}(\mathcal{H})$ be an operator with discrete spectrum, say $\sigma(T) = \{\lambda_n : n \in \mathbb{N}\}$, such that

1. $(\lambda_n)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$,
2. each λ_n is an eigenvalue of T, with eigenvector given by $e_n \in \mathcal{H}$, and
3. $\{e_n : n \in \mathbb{N}\}$ is an orthonormal basis for $\mathcal{H}$.

This is the case, for example, when T is compact and normal. We now define a map $U : \ell^2(\mathbb{N}) \to \mathcal{H}$ as follows

$$U(x_n)_n = \sum_{n\in\mathbb{N}} x_n e_n$$

One checks that U is well defined linear map. Further, for any $\xi \in \mathcal{H}$ we have

$$\langle U(x_n)_n, \xi\rangle = \sum_{n\in\mathbb{N}} x_n \langle e_n, \xi\rangle = \sum_{n\in\mathbb{N}} x_n \overline{\langle \xi, e_n\rangle} = \langle (x_n)_n, (\langle \xi, e_n\rangle)_n\rangle_{\ell^2}$$

Thus, we have $U^* : \mathcal{H} \to \ell^2(\mathbb{N})$ is given by

$$U^*\xi = (\langle \xi, e_n\rangle)_n$$

Now, using Parseval's identity we get

$$UU^*\xi = \sum_{n\in\mathbb{N}} \langle \xi, e_n\rangle e_n = \xi,$$

which gives $UU^* = \mathrm{id}_{\mathcal{H}}$. Also, since $(e_n)_n$ is orthonormal, we have

$$U^*U(x_n)_n = (\langle \sum_{k\in\mathbb{N}} x_k e_k, e_n\rangle)_n = (\sum_{k\in\mathbb{N}} x_k \langle e_k, e_n\rangle)_n = (x_n)_n,$$

and therefore $U^*U = \mathrm{id}_{\ell^2(\mathbb{N})}$. Next, we let $M := U^*TU : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$. Since $Te_k = \lambda_k e_k$, it follows that

$$M(x_n)_n = U^* \sum_{n\in\mathbb{N}} x_n \lambda_n e_n = \sum_{n\in\mathbb{N}} x_n \lambda_n (\langle e_n, e_k\rangle)_k = (x_n\lambda_n)_n$$

This is saying that M is in fact a "multiplication operator" by $(\lambda_n)_n$, and moreover this forces T to be a normal operator. Indeed,

$$\langle M(x_n)_n, (y_n)_n\rangle = \langle (x_n\lambda_n)_n, (y_n)_n\rangle = \sum_{n\in\mathbb{N}} x_n \lambda_n \overline{y_n} = \sum_{n\in\mathbb{N}} x_n \overline{\overline{\lambda_n} y_n} = \langle x_n, \overline{\lambda_n} y_n\rangle.$$

This gives, $M^*(y_n)_n = \overline{\lambda_n} y_n$ and therefore

$$M^*M(x_n)_n = (|\lambda_n|^2 x_n) = MM^*(x_n)_n,$$

so M is normal. Thus, since $T = UMU^*$, we have

$$TT^* = UMU^*UM^*U^* = UM^*MU^* = (UMU^*)^*(UM^*U^*) = T^*T$$

▼

As a remark from the previous example, we notice that on the infinite dimensional case, we are replacing the diagonal matrix Λ with a "multiplication operator" M. We say more about what a multiplication operator is below.

Basic Definitions

Definition. (Multiplication Operator) Let $(\Omega, \mathfrak{M}, \mu)$ be a σ-finite measure space and $f \in L^\infty(\Omega, \mu)$. We define an operator M_f on $L^2(\Omega, \mu)$ by

$$M_f(g)(\omega) := f(\omega)g(\omega)$$

for any $g \in L^2(\Omega, \mu)$ and $\omega \in \Omega$. The operator M_f is called **multiplication by** f. ▲

Properties. *For any $f, h \in L^\infty(\Omega, \mu)$ we have*

- $M_f \in \mathcal{B}(L^2(\Omega, \mu))$
- $\|M_f\| = \|f\|_\infty$
- $M_{f+h} = M_f + M_h$
- $\sigma(M_f) = \overline{f(\Omega)}$

Definition. Let $\mathcal{H}$ be a separable Hilbert space. An operator $T \in \mathcal{B}(\mathcal{H})$ is said to be **diagonalizable** if there exist

1. a σ-finite measure space $(\omega, \mathfrak{M}, \mu)$ such that $L^2(\Omega, \mu)$ is separable,
2. a function $f \in L^\infty(\Omega, \mu)$, and
3. a unitary operator $U : L^2(\Omega, \mu) \to \mathcal{H}$ such that $T = UM_fU^*$.

▲

We are now ready to state the most general version of the spectral theorem:

Theorem. *(The Spectral Theorem) Let $\mathcal{H}$ be a separable Hilbert space and $T \in \mathcal{B}(\mathcal{H})$ a normal operator. Then T is diagonalizable.*

Before sketching the proof, we give an immediate corollary and an example.

Corollary. *Let $\mathcal{H}$ be a separable Hilbert space and $T \in \mathcal{B}(\mathcal{H})$ a normal operator. Then $\sigma(T) = \overline{f(\Omega)}$, where f is the function in $L^\infty(\Omega, \mu)$ that diagonalizes T.*

Proof. Since $T = UM_fU^*$, we have

$$\sigma(T) = \sigma(UM_fU^*) = \sigma(M_f) = \overline{f(\Omega)}$$

■

Example. Let $\mathcal{H} = \ell^2(\mathbb{Z})$ and $T : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ be given by

$$T(x_n)_n = (x_{n-1} + x_{n+1})_n$$

It's easily checked that $T \in \mathcal{B}(\ell^2(\mathbb{Z}))$. Further, one checks that T is self adjoint (i.e. $T^* = T$) and therefore T is normal. Hence, T is diagonalizable.

We now find all the ingredients that make T diagonalizable. Let $\Omega = S^1$ with μ normalized Lebesgue measure on S^1 and σ-algebra $\mathfrak{B}(S^1)$ given by all the Borel sets on S^1. We know that $(S^1, \mathfrak{B}(S^1), \mu)$ is a finite measure space, in fact $\mu(S^1) = 1$. Next, we define functions $u_n : S^1 \to S^1$ by $u_n(e^{i\theta}) = e^{in\theta}$. Then, $\{u_n : n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(S^1, \mu)$. We let $U : L^2(S^1, \mu) \to \ell^2(\mathbb{Z})$ be given by

$$Ug := \left(\int_{S^1} g\overline{u_n} d\mu \right)_n .$$

We check that U is unitary. Indeed, first notice that

$$\langle Ug, (x_n)_n \rangle = \sum_{n\in\mathbb{Z}} \left(\int_{S^1} g\overline{u_n}\ d\mu \right) \overline{x_n} = \int_{S^1} g \left(\overline{\sum_{n\in\mathbb{Z}} u_n x_n} \right) d\mu = \langle g, \sum_{n\in\mathbb{Z}} u_n x_n \rangle$$

whence

$$U^*(x_n)_n = \sum_{n\in\mathbb{Z}} u_n x_n$$

Therefore, by orthonormality of $\{u_n : n \in \mathbb{Z}\}$ and since $\mu(S^1) = 1$ we have

$$UU^*(x_n)_n = \left(\int_{S^1} \left(\sum_{n\in\mathbb{Z}} u_n x_n \right) \overline{u_k} d\mu \right)_k = (x_n)_n,$$

also, by Parseval we get

$$U^*Ug = \sum_{n\in\mathbb{Z}} u_n \left(\int_{S^1} g\overline{u_n} d\mu \right) = \sum_{n\in\mathbb{Z}} u_n \langle g, u_n \rangle = g$$

This gives that U is a unitary operator. Next, we define $f : S^1 \to \mathbb{R}$ by

$$f(e^{i\theta}) = 2\cos(\theta)$$

Since $|f| \leq 2$, it follows that $f \in L^\infty(S^1, \mu)$. Furthermore, notice that $f = u_{-1} + u_1$ (this is because $e^{-i\theta} + e^{i\theta} = 2\cos(\theta)$). Then, we compute

$$\begin{aligned} UM_f g = U(u_{-1}g + u_1 g) &= \left(\int_{S^1} (u_{-1}g + u_1 g)\overline{u_n} d\mu \right)_n \\ &= \left(\int_{S^1} (u_{-1}\overline{u_n} + u_1\overline{u_n}) g d\mu \right)_n \\ &= \left(\int_{S^1} (\overline{u_{n+1}} + \overline{u_{n-1}}) g d\mu \right)_n \\ &= \left(\int_{S^1} g\overline{u_{n+1}}\ d\mu + \int_{S^1} g\overline{u_{n-1}}\ d\mu \right)_n \\ &= T \left(\int_{S^1} g\overline{u_n}\ d\mu \right)_n \\ &= TUg \end{aligned}$$

That is, $UM_f = TU$, and therefore $T = UM_f U^*$, as we wanted to show. ▼

Remark. Notice that the T from the example above is not compact. In fact, $\sigma(T) = \overline{f(S^1)} = [-2, 2]$, so T can't have discrete spectrum consisting of eigenvalues. ▼

We now give a sketch of the proof for the spectral theorem:

Theorem. *(The Spectral Theorem) Let $\mathcal{H}$ be a separable Hilbert space and $T \in \mathcal{B}(\mathcal{H})$ a normal operator. Then T is diagonalizable.*

Sketch of Proof. Let $\mathcal{A} := C^*(\{I, T\})$, the C^* algebra generated by T and the identity.

Step 1: By separability of $\mathcal{H}$, we can write $\mathcal{H} = \bigoplus_{n \in \mathbb{J}} \mathcal{H}_n$, where $\mathbb{J} \subseteq \mathbb{N}$ and for each $n \in \mathbb{J}$, there is $\xi_n \in \mathcal{H}_n$ such that

$$\overline{\mathcal{A}\xi_n} = \overline{\{S\xi_n : S \in \mathcal{A}\}} = \mathcal{H}_n$$

Step 2: Fix $n \in \mathbb{J}$. Define $T_n := T \mid_{\mathcal{H}_n}$. Since T is normal, T_n is also normal. Let $\Omega_n = \sigma(T_n)$. Define $\varphi_n : C(\Omega_n) \to \mathbb{C}$ by

$$\varphi_n(f) := \langle f(T_n)\xi_n, \xi_n \rangle,$$

where $f \mapsto f(T_n)$ is the functional calculus for T_n. One checks that φ_n is a positive linear functional, and therefore, by the Riesz-Markov representation theorem there is a positive Borel measure μ_n on Ω_n such that

$$\varphi_n(f) = \int_{\Omega_n} f d\mu_n.$$

Further, since Ω_n is compact, we must have $\mu_n(\Omega_n) < \infty$. Standard computations give $\|f(T_n)\xi_n\| = \|f\|_{L^2(\Omega_n, \mu_n)}$ and therefore $U_n f := f(T_n)\xi_n$ defines a unitary operator $U_n : C(\Omega_n) \to \mathcal{A}_n \xi_n$ where $\mathcal{A}_n := C^*(\{I, T_n\})$. This operator extends by density to $U_n : L^2(\Omega_n, \mu_n) \to \mathcal{H}_n$ and satisfies $U_n M_f g = f(T_n) U_n g$ for all $f \in C(\Omega_n)$, for all $g \in L^2(\Omega_n, \mu_n)$. In particular, if $\iota_n : \Omega_n \to \mathbb{C}$ is the canonical inclusion, if follows that $U_n M_{\iota_n} = T_n U_n$. This gives that T_n is diagonalizable.

Step 3: Finally, since each T_n is diagonalizable on $\mathcal{H}_n$ by *Step 2*, one checks that this implies that T is diagonalizable on $\mathcal{H}$. "□"