Review of Spectral Theory

Definition 1 Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathcal{H})$.

(a) A is called *self–adjoint* if $A = A^*$.

(b) A is called *unitary* if $A^*A = AA^* = \mathbb{1}$. Equivalently, A is unitary if it is bijective (i.e. 1–1 and onto) and preserves inner products.

(c) A is called *normal* if $A^*A = AA^*$. That is, if A commutes with its adjoint.

(d) Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. A linear operator $C : \mathcal{X} \to \mathcal{Y}$ is said to be compact if for each bounded sequence $\{x_i\}_{i\in\mathbb{N}} \subset \mathcal{X}$, there is a subsequence of $\{Cx_i\}_{i\in\mathbb{N}}$ that is convergent.

Before we start on spectral theory in infinite dimensions, here is a series of remarks that reviews the spectral theory of matrices.

Remark 2 Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathcal{H})$. In this remark, we assume that $\mathcal{H}$ is finite dimensional, so that A is multiplication by a matrix, that we also denote A. So let A be an $n \times n$ matrix.

(a) By definition, λ is an eigenvalue of A with eigenvector $\mathbf{x} \neq \mathbf{0}$ if $A\mathbf{x} = \lambda\mathbf{x}$. So

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } A &\iff (\lambda\mathbb{1} - A) \text{ has a nontrivial kernel} \\
&\iff (\lambda\mathbb{1} - A) \text{ does not have an inverse matrix} \\
&\qquad \text{since } \dim\operatorname{kernel}(\lambda\mathbb{1} - A) + \dim\operatorname{range}(\lambda\mathbb{1} - A) = n \\
&\iff \det(\lambda\mathbb{1} - A) = 0
\end{aligned}$$

We are not going to be able to use the $\det(\lambda\mathbb{1} - A) = 0$ test when the dimension is infinite, because $\det(\lambda\mathbb{1} - A)$ will typically not be defined. For example, if A is the $n \times n$ zero matrix, $\det(\lambda\mathbb{1} - A) = \det(\lambda\mathbb{1}) = \lambda^n$. As $n \to \infty$ this diverges when $|\lambda| \geq 1$, $\lambda \neq 1$ and converges to 0 when $|\lambda| < 1$.

(b) Suppose that A has n linearly independent eigenvectors (which is always case if A has n distinct eigenvalues, or if $A = A^*$ or if $A^*A = \mathbb{1}$). Call the eigenvectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$ and the corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$. Write $V = \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}$ (that is, the columns of V are eigenvectors of A) and denote by D the diagonal matrix whose diagonal entries are the λ_j's. Then

$$\begin{aligned} AV &= \big[A\mathbf{x}_1 \ \cdots \ A\mathbf{x}_n\big] = \big[\lambda\mathbf{x}_1 \ \cdots \ \lambda_n\mathbf{x}_n\big] = VD \\ \text{or} \quad V^{-1}AV &= D \\ \text{or} \quad V^{-1}AV \begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix} &= \begin{bmatrix}\lambda_1\alpha_1\\ \lambda_2\alpha_2\\ \vdots\\ \lambda_n\alpha_n\end{bmatrix} \end{aligned}$$

That is $V^{-1}AV$ is a multiplication operator. Replacing A by $V^{-1}AV$ amounts to a change of basis. So if A has n independent eigenvectors, we can pick a basis for $\mathcal{H}$ with respect to which A is a multiplication operator.

(c) The infinite dimensional case will be much more interesting and complicated. We can see this even for operators that are already multiplication operators. For example, let $a(x)$ be a bounded measurable function on $[0, 1]$ and consider the operator

$$(A\varphi)(x) = a(x)\,\varphi(x)$$

on $L^2((0, 1))$.

It is certainly possible for A to have eigenvalues. For example if, for some $\lambda \in \mathbb{C}$, we have $\mu\big(a^{-1}(\{\lambda\})\big) > 0$ (that is, $a(x)$ takes the value λ on a set of strictly positive measure), then

$$a(x)\chi_{a^{-1}(\{\lambda\})}(x) = \lambda\,\chi_{a^{-1}(\{\lambda\})}(x)$$

and, as $\chi_{a^{-1}(\{\lambda\})}$ is not the zero vector, λ is an eigenvalue with eigenvector $\chi_{a^{-1}(\{\lambda\})}$.

On the other hand it is also possible for A to have absolutely no eigenvalues. For example, if $a(x) = x$, then $a(x)$ takes all real values between 0 and 1, but, for any $\lambda \in \mathbb{C}$,

$$\begin{aligned} A\varphi = \lambda\varphi \iff & \quad a(x)\,\varphi(x) = \lambda\,\varphi(x) && \text{for almost every } x \in [0, 1] \\ \iff & \quad \big(x - \lambda\big)\,\varphi(x) = 0 && \text{for almost every } x \in [0, 1] \\ \iff & \quad \varphi(x) = 0 && \text{for almost every } x \in [0, 1] \end{aligned}$$

since $x - \lambda$ is nonzero everywhere except at the single point $x = \lambda$.

Remark 3 Again, let A be an $n \times n$ matrix.

(a) If A is self–adjoint, we have:

(i) All eigenvalues of A are real, since

$$\mathbf{0} \neq \mathbf{x} \in \mathcal{H},\ A\mathbf{x} = \lambda\mathbf{x} \implies \lambda\,\langle\mathbf{x}, \mathbf{x}\rangle = \langle\mathbf{x}, A\mathbf{x}\rangle = \langle A\mathbf{x}, \mathbf{x}\rangle = \overline{\lambda}\,\langle\mathbf{x}, \mathbf{x}\rangle$$

(ii) Eigenvectors of A that correspond to different eigenvalues are perpendicular, since

$$\mathbf{0} \ne \mathbf{x} \in \mathcal{H},\ A\mathbf{x} = \lambda\mathbf{x},\quad \mathbf{0} \ne \mathbf{y} \in \mathcal{H},\ A\mathbf{y} = \mu\mathbf{y},\quad \lambda \ne \mu$$
$$\implies \lambda \langle \mathbf{x}, \mathbf{y} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle = \mu \langle \mathbf{x}, \mathbf{y} \rangle$$

(iii) There is an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of A. See the notes "Families of Commuting Normal Matrices". Call the basis vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$. Write $U = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix}$ (that is, the columns of U are the eigenvectors of A) and denote by D the diagonal matrix whose diagonal entries are the corresponding eigenvalues. Then U is unitary (the condition that $U^*U = \mathbb{1}$ is the same as the condition that the columns are orthonormal) and, as we saw in (b) above,

$$AU = UD \quad \text{or} \quad U^{-1}AU = D \quad \text{or} \quad U^*AU = D$$

(b) This and similar arguments give that

$$\begin{aligned} A \text{ is self–adjoint} &\iff \mathcal{H} \text{ has an orthonormal basis of eigenvectors of } A \text{ and} \\ &\ \text{all eigenvalues of } A \text{ are real} \\ A \text{ is unitary} &\iff \mathcal{H} \text{ has an orthonormal basis of eigenvectors of } A \text{ and} \\ &\ \text{all eigenvalues } \lambda \text{ of } A \text{ obey } |\lambda| = 1 \\ A \text{ is normal} &\iff \mathcal{H} \text{ has an orthonormal basis of eigenvectors of } A \end{aligned}$$

(c) When $\mathcal{H}$ is finite dimensional, A is normal if and only if there is a unitary matrix U and a diagonal matrix D such that $U^*AU = D$. That is, A is diagonalizable by a unitary matrix. The corresponding statement when $\mathcal{H}$ is infinite dimensional, is that A is normal if and only if there is a unitary operator U such that U^*AU is a multiplication operator. We shall prove this.

(d) If an operator is either self–adjoint or unitary, it is also normal.

(e) An operator A is normal if and only if it can be written in the form $A = B + iC$ with B and C self–adjoint and commuting. (Take $B = \frac{1}{2}(A + A^*)$ and $C = \frac{1}{2i}(A - A^*)$.)

Remark 4 In quantum mechanics, "physical observables" tend to be self–adjoint operators — energy, momentum, etc are real quantities. In quantum mechanics, time evolution is by a unitary operator. "Total probability is preserved."

Remark 5 When $\mathcal{H}$ is finite dimensional, all operators are compact. Even when $\mathcal{H}$ is infinite dimensional, compact operators behave a lot like finite dimensional matrices. See Theorem 20. They tend to be easier to work with than other operators.

Remark 6 In the finite dimensional case, we say that A is diagonalizable if there is an invertible (but not necessarily unitary) matrix V such that VAV^{-1} is diagonal. Then $\mathcal{H}$ has a basis of eigenvectors of A, but the basis vectors need not be mutually perpendicular.

A natural extension of this finite dimensional definition to infinite dimensions would be that A is diagonalizable if there is a bounded linear bijection $V : \mathcal{H} \to \mathcal{H}' = L^2(\mathcal{M}, \mu)$ (then it has a bounded inverse by the inverse mapping theorem) such that VAV^{-1} is a multiplication operator. Again, V need not be unitary. But we can make V unitary by changing the inner product on $\mathcal{H}$ (thereby changing the meaning of "orthogonal"). Define

$$\langle \mathbf{x}, \mathbf{y} \rangle'_{\mathcal{H}} = \langle V\mathbf{x}, V\mathbf{y} \rangle_{\mathcal{H}'}$$

Then V is unitary as a map from $\big(\mathcal{H}, \langle \cdot\,, \cdot \rangle'_{\mathcal{H}}\big)$ to $\mathcal{H}'$. Here $\big(\mathcal{H}, \langle \cdot\,, \cdot \rangle'_{\mathcal{H}}\big)$ is the vector space $\mathcal{H}$ equipped with the inner product $\langle \cdot\,, \cdot \rangle'_{\mathcal{H}}$ instead of the inner product $\langle \cdot\,, \cdot \rangle_{\mathcal{H}}$. Changing the inner product in this way changes the lengths of vectors in $\mathcal{H}$ and also the angles between vectors in $\mathcal{H}$, but does not change the topology since

$$\|V^{-1}\|^{-1} \|\mathbf{x}\|_{\mathcal{H}} \le \|\mathbf{x}\|'_{\mathcal{H}} \le \|V\| \, \|\mathbf{x}\|_{\mathcal{H}}$$

As A is diagonalizable by a unitary operator if and only if it is normal, we have that A is diagonalizable if and only if A is normal with respect to some inner product on $\mathcal{H}$ that gives a topology equivalent to the original topology. (All such inner products are of the form $= \langle V\mathbf{x}, V\mathbf{y} \rangle_{\mathcal{H}}$ for some bijection V on $\mathcal{H}$ that is bounded with bounded inverse. See Corollary 35 in the notes "Review of Hilbert and Banach Spaces".)

Definition 7 Let $\mathcal{B}$ be a Banach space and $T \in \mathcal{L}(\mathcal{B})$.

(a) The *resolvent set*, $\rho(T)$, of T is the set of all complex numbers λ such that $\lambda \mathbb{1} - T$ is a bijection with bounded inverse.

(b) The *resolvent* of T at $\lambda \in \rho(T) \subset \mathbb{C}$ is $R_\lambda(T) = \big(\lambda \mathbb{1} - T\big)^{-1}$.

(c) The *spectrum* of T is $\sigma(T) = \mathbb{C} \setminus \rho(T)$.

(d) The complex number λ is in the *point spectrum*, $\sigma_p(T)$, of T if $\lambda \mathbb{1} - T$ is not injective. That is, if there is a nonzero vector $\mathbf{x} \in \mathcal{H}$ such that $T\mathbf{x} = \lambda\mathbf{x}$. Then $\mathbf{x}$ is said to be an eigenvector of T with eigenvalue λ.

(e) The complex number λ is in the *residual spectrum*, $\sigma_r(T)$, of T if $\lambda \mathbb{1} - T$ is injective but the range of T is not dense in $\mathcal{B}$.

(f) The complex number λ is in the *continuous spectrum*, $\sigma_c(T)$, of T if $\lambda \mathbb{1} - T$ is injective and the range of T is dense in, but not all of, $\mathcal{B}$.

Remark 8 There is no universally accepted definition of "continuous spectrum". I have just chosen the simplest contender.

Remark 9

(a) In finite dimensions there is no residual spectrum, because the dimensions of the kernel and the range of an $n \times n$ matrix always add to exactly n. So if the dimension of the range is strictly less than n, then the dimension of the kernel is necessarily at tleast 1.

(b) In the finite dimensions there is no continuous spectrum, just because the range of any $n \times n$ matrix is always closed.

Remark 10 Let M be an $n \times n$ matrix. By definition, M is diagonalizable if there is an invertible matrix V and a diagonal matrix D such that $V^{-1}MV = D$. In this case the columns of V are eigenvectors of M that form a basis for $\mathbb{C}^n$, the diagonal elements of D are the eigenvalues of M and $\sigma(M) = \sigma_p(M)$ is the set of all eigenvalues of M.

If M is not diagonalizable, there still exists an invertible matrix V such that $V^{-1}MV$ is in Jordan form. This means that it is of the form

$$\begin{bmatrix} J_1 & 0 & \cdots & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & \cdots & 0 & J_m \end{bmatrix}$$

with each diagonal block being a Jordan block. A 4×4 Jordan block is of the form

$$B_{\lambda,4} = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

Observe that $\lambda\mathbb{1} - B_{\lambda,4}$ has range $\big\{ (z_1, z_2, z_3, 0) \bigm| z_1, z_2, z_3 \in \mathbb{C} \big\}$ which is not dense. In this finite dimensional world, $\lambda\mathbb{1} - B_{\lambda,4}$ must also have a kernel (since the dimension of the range plus the dimension of the kernel must be the dimension of the world). That is, λ must be an eigenvalue. For an infinite dimensional world that is no longer the case. The existence of residual spectrum signals the failure of diagonalizability.

Example 11

Multiplication Operators: Let

- $(X, \mathcal{M}, \mu)$ be a semifinite measure space,

- $1 \le p \le \infty$ and
- $a : X \to \mathbb{C}$ be a bounded measurable function on X.

Define the bounded operator $A : L^p(X, \mathcal{M}, \mu) \to L^p(X, \mathcal{M}, \mu)$ by

$$(A\varphi)(x) = a(x)\varphi(x)$$

Then

$$\begin{aligned}
\rho(A) &= \big\{\ \lambda \in \mathbb{C}\ \big|\ \exists\ \varepsilon > 0 \text{ such that } |\lambda - a(x)| \ge \varepsilon \text{ a.e. } \big\} \\
\sigma_p(A) &= \big\{\ \lambda \in \mathbb{C}\ \big|\ \mu\big(\{x \in X \,|\, a(x) = \lambda\}\big) > 0\ \big\} \\
\sigma_r(A) &= \begin{cases} \emptyset & \text{if } 1 \le p < \infty \\ \big\{\ \lambda \in \mathbb{C}\ \big|\ \nexists\ \varepsilon > 0 \text{ such that } |\lambda - a(x)| \ge \varepsilon \text{ a.e. } \big\} \setminus \sigma_p(A) & \text{if } p = \infty \end{cases}
\end{aligned}$$

Shift Operators: Define the right and left shift operators acting on ℓ^2 by

$$\begin{aligned}
L(\alpha_1, \alpha_2, \alpha_3, \cdots) &= (\alpha_2, \alpha_3, \cdots) \\
R(\alpha_1, \alpha_2, \alpha_3, \cdots) &= (0, \alpha_1, \alpha_2, \alpha_3, \cdots)
\end{aligned}$$

Then

$$\begin{aligned}
\rho(L) &= \big\{\ \lambda \in \mathbb{C}\ \big|\ |\lambda| > 1\ \big\} & \rho(R) &= \big\{\ \lambda \in \mathbb{C}\ \big|\ |\lambda| > 1\ \big\} \\
\sigma_p(L) &= \big\{\ \lambda \in \mathbb{C}\ \big|\ |\lambda| < 1\ \big\} & \sigma_p(R) &= \emptyset \\
\sigma_r(L) &= \emptyset & \sigma_r(R) &= \big\{\ \lambda \in \mathbb{C}\ \big|\ |\lambda| < 1\ \big\}
\end{aligned}$$

See the notes "Spectral Theory Examples" for derivations and other examples.

Definition 12 Let $\mathcal{X}$ be a Banach space and $\mathcal{D}$ an open subset of $\mathbb{C}$. A function $\mathbf{x} : \mathcal{D} \to \mathcal{X}$ is analytic at $z_0 \in \mathcal{D}$ if

$$\lim_{z \to z_0} \tfrac{\mathbf{x}(z) - \mathbf{x}(z_0)}{z - z_0}$$

exists.

Lemma 13 *Let B be a bounded linear operator on the Banach space $\mathcal{B}$. Assume that B has a bounded inverse and that C is a bounded operator on $\mathcal{B}$ with $\|C\| < \big\|B^{-1}\big\|^{-1}$. Then $B + C$ is 1–1 and onto and has a bounded inverse and*

$$(B + C)^{-1} = \sum_{n=0}^{\infty} \big(-B^{-1}C\big)^n B^{-1} = B^{-1} - B^{-1}CB^{-1} + B^{-1}CB^{-1}CB^{-1} - \cdots$$

with convergence in norm. Furthermore

$$\big\|(B + C)^{-1}\big\| \le \tfrac{\|B^{-1}\|}{1 - \|B^{-1}\|\,\|C\|} \qquad \big\|(B + C)^{-1} - B^{-1}\big\| \le \tfrac{\|B^{-1}\|^2\,\|C\|}{1 - \|B^{-1}\|\,\|C\|}$$

Theorem 14 *Let $\mathcal{B}$ be a Banach space and $T \in \mathcal{L}(\mathcal{B})$. Then*

(a) $\rho(T)$ is an open subset of $\mathbb{C}$. Furthermore, if $\lambda \in \rho(T)$, then

$$\lim_{\mu \to \lambda} \big\| R_\mu(T) - R_\lambda(T) \big\| = 0$$

(b) $R_\lambda(T)$ is an analytic $\mathcal{L}(\mathcal{B})$–valued function of λ on $\rho(T)$.

(c) First resolvent formula: *If $\lambda, \mu \in \rho(T)$, then $R_\lambda(T)$ and $R_\mu(T)$ commute and*

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda) R_\mu(T) R_\lambda(T)$$

Second resolvent formula: *If $S \in \mathcal{L}(\mathcal{B})$ and $\lambda \in \rho(S) \cap \rho(T)$, then*

$$R_\lambda(S) - R_\lambda(T) = R_\lambda(S)\,(S - T)\,R_\lambda(T) = R_\lambda(T)\,(S - T)\,R_\lambda(S)$$

(d) If $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$.

(e) $\sigma(T) \neq \emptyset$.

Lemma 15 *Let B be a bounded linear operator on the Banach space $\mathcal{B}$. If $P(z)$ is a polynomial and $\lambda \in \sigma(B)$ then $P(\lambda) \in \sigma\big(P(B)\big)$.*

Theorem 16 *Let $\mathcal{B}$ be a Banach space and $T \in \mathcal{L}(\mathcal{B})$. The spectral radius of T is defined to be*

$$r(T) = \sup_{\lambda \in \sigma(T)} |\lambda|$$

We have

$$r(T) = \lim_{n \to \infty} \big\| T^n \big\|^{\frac{1}{n}}$$

If $\mathcal{B}$ is a Hilbert space and T is self–adjoint, then

$$r(T) = \|T\|$$

Theorem 17 *Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathcal{H})$.*
(a) $\sigma(A^*) = \overline{\sigma(A)}$.
(b) $\lambda \in \sigma_r(A) \implies \overline{\lambda} \in \sigma_p(A^*)$
$\lambda \in \sigma_p(A) \implies \overline{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$

Lemma 18 *Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{L}(\mathcal{H})$ be normal.*

(a) If φ is an eigenvector of A of eigenvalue λ, then φ is an eigenvector of A^* of eigenvalue $\bar{\lambda}$.

(b) Eigenvectors of A with different eigenvalues are orthogonal.

(c) A has no residual spectrum.

(d) If $A = A^*$, then $\sigma(A) \subset \mathbb{R}$.

(e) $\lambda \in \sigma(A)$ if and only if, for each $\varepsilon > 0$, there exists a $\varphi \in \mathcal{H}$ with $\|\varphi\| = 1$ and $\|(\lambda \mathbb{1} - A)\varphi\| < \varepsilon$.

Theorem 19 *Let $\mathcal{H}$ and $\mathcal{H}'$ be Hilbert spaces. Let $U \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ be bijective and $A \in \mathcal{L}(\mathcal{H})$ and set $A' = UAU^{-1} \in \mathcal{L}(\mathcal{H}')$. Then*

$$\rho(A') = \rho(A) \qquad \sigma(A') = \sigma(A) \qquad \sigma_p(A') = \sigma_p(A) \qquad \sigma_r(A') = \sigma_r(A)$$

Theorem 20 (The Spectrum of Compact Operators) *Let $C : \mathcal{X} \to \mathcal{X}$ be a compact operator on the Banach space $\mathcal{X}$. The spectrum of C consists of at most countably many points. For any $\varepsilon > 0$, $\{\ \lambda \in \sigma(C)\ \big|\ |\lambda| > \varepsilon\ \}$ is finite. If $0 \neq \lambda \in \sigma(C)$, then λ is an eigenvalue of C of finite multiplicity.*

Theorem 21 (Spectral Theorem - Multiplication Operator Version)
Let A be a bounded self–adjoint operator on a Hilbert space $\mathcal{H}$. There exist

- *a measure space $\langle M, \Sigma, \mu \rangle$,*
- *a bounded measurable function $a : M \to \mathbb{R}$, and*
- *a unitary operator $U : \mathcal{H} \to L^2(M, \Sigma, \mu)$*

such that

$$\big(UAU^{-1}\varphi\big)(m) = a(m)\,\varphi(m)$$

for all $\varphi \in L^2(M, \Sigma, \mu)$. If $\mathcal{H}$ is separable, μ can be chosen to be a finite measure.

Example 22 Let A be a self–adjoint, compact operator and let $\{\varphi_n\}_{n \in \mathcal{I}}$ be a complete orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of A. Denote by λ_n the eigenvalue of A for the eigenvector φ_n.

Think of $\ell^2(\mathcal{I})$ as L^2 of the measure space $\mathcal{I}$, equipped with the counting measure. So, think of an element of $\ell^2(\mathcal{I})$ as a function on $\mathcal{I}$ rather than a sequence. Define the unitary operator $U : \mathcal{H} \to \ell^2(\mathcal{I})$ by

$$\Big(U\big(\sum_{n \in \mathcal{I}} x_n \varphi_n\big)\Big)(m) = x_m$$

The inverse operator $U^{-1} = U^* : \ell^2(\mathcal{I}) \to \mathcal{H}$ is given by

$$U^{-1}v = \sum_{n\in\mathcal{I}} v(n)\varphi_n$$

For each $m \in \mathbb{N}$, denote by e_m the element of $\ell^2(\mathcal{I})$ all of whose components are zero except for the $m^{\rm th}$, which is 1. That is

$$e_m(n) = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

Observe that $U\varphi_m = e_m$, for each $m \in \mathcal{I}$. Also define the function $a : \mathbb{N} \to \mathbb{R}$ by $a(m) = \lambda_m$. Then, for each $v \in \ell^2(\mathcal{I})$ and each $m \in \mathcal{I}$,

$$\begin{aligned} (UAU^{-1}v)(m) &= \big\langle e_m, UAU^{-1}v\big\rangle_{\ell_2} = \big\langle U^{-1}e_m, AU^{-1}v\big\rangle_{\mathcal{H}} = \big\langle \varphi_m, AU^{-1}v\big\rangle_{\mathcal{H}} \\ &= \big\langle A\varphi_m, U^{-1}v\big\rangle_{\mathcal{H}} \quad = \lambda_m \left\langle \varphi_m, U^{-1}v\right\rangle_{\mathcal{H}} \quad = \lambda_m \left\langle U\varphi_m, v\right\rangle_{\ell_2} \\ &= \lambda_m \left\langle e_m, v\right\rangle_{\ell_2} \qquad\quad = a(m)v(m) \end{aligned}$$

If, as will often be the case, $\mathcal{I} = \mathbb{N}$, the counting measure on $\mathcal{I}$ is not finite. But it is easy to rework the above construction so as to use a finite measure space. Define the measure μ on $\mathbb{N}$ by $\mu\big(\{m\}\big) = \frac{1}{2^m}$. Then define $U : \mathcal{H} \to L^2(\mathbb{N}, \mu)$ by $U\Big(\sum_{n\in\mathcal{I}} x_n\varphi_n\Big)(m) = 2^{m/2}x_m$ and the function a by $a(m) = \lambda_m$ again.

Theorem 23 (Spectral Theorem - Multiplication Operator, with Multiplicity) *Let A be a bounded self–adjoint operator on a separable Hilbert space $\mathcal{H}$. There exist $N \in \mathbb{N} \cup \{\infty\}$ and measures μ_n, $1 \le n \le N$, on the spectrum $\sigma(A) \subset \mathbb{R}$, of A, and a unitary operator $U : \mathcal{H} \to \oplus_{n=1}^N L^2(\sigma(A), \mu_n)$ such that*

$$\big(UAU^{-1}\varphi\big)_n(\lambda) = \lambda\,\varphi_n(\lambda)$$

for all $\varphi \in \oplus_{n=1}^N L^2(\sigma(A), \mu_n)$.

Example 24 Let A be a self–adjoint, compact operator and let $\big\{\varphi_n\big\}_{n\in\mathcal{I}}$ be a complete orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of A. Denote by λ_n the eigenvalue of A for the eigenvector φ_n. Denote by $1 \le N \le \infty$ the supremum of the multiplicities of the eigenvalues of A. Define, for each $1 \le n \le N$, the Borel measure μ_n on $\mathbb{R}$ by

$$\mu_n(B) = \sum_{n\in\mathcal{I}} \begin{cases} 1 & \text{if } \lambda_n \in B \text{ and } \lambda_n \text{ has multiplicity at least } n \\ 0 & \text{otherwise} \end{cases}$$

Then $\oplus_{n=1}^{N} L^2(\mathbb{N}, \mu_n)$ is the set of all N–vector valued functions on $\mathbb{R}$ with

$$\|\vec{\psi}\|^2 = \sum_{n=1}^{N} \int d\mu_n(x)\, \big|\psi_n(x)\big|^2$$

Define the unitary operator $U\mathcal{H} \to \oplus_{n=1}^{N} L^2(\mathbb{N}, \mu_n)$ by

$$U\Big(\sum_{m=1}^{\infty} x_m\varphi_m\Big)_n(\lambda) = \sum_{m=1}^{\infty} x_m \begin{cases} 1 & \text{if } \varphi_m \text{ is the } n^{\rm th} \text{ eigenvector of eigenvalue } \lambda \\ 0 & \text{otherwise} \end{cases}$$

Theorem 25 (Spectral Theorem - Commuting Operators Version)
Let $A_1, \cdots, A_n$ be a finite set of commuting, bounded, self–adjoint operators on a Hilbert space $\mathcal{H}$. There exist

- *a measure space $\langle M, \mu\rangle$,*
- *bounded measurable functions $a_\ell : M \to \mathbb{R}$, $1 \le \ell \le n$, and*
- *a unitary operator $U : \mathcal{H} \to L^2(M, \mu)$*

such that, for each $1 \le \ell \le n$,

$$\big(UA_\ell U^{-1}\varphi\big)(m) = a_\ell(m)\,\varphi(m)$$

for all $\varphi \in L^2(M, \mu)$. If $\mathcal{H}$ is separable, μ can be chosen to be a finite measure.

Corollary 26 *Let A be a bounded normal operator on a Hilbert space $\mathcal{H}$. There exist a measure space $\langle M, \mu\rangle$, a bounded measurable function $a : M \to \mathbb{C}$, and a unitary operator $U : \mathcal{H} \to L^2(M, \mu)$ such that*

$$\big(UAU^{-1}\varphi\big)(m) = a(m)\,\varphi(m)$$

for all $\varphi \in L^2(M, \mu)$.

Theorem 27 (Spectral Theorem - Functional Calculus Version)
Let A be a bounded self–adjoint operator on a Hilbert space $\mathcal{H}$. Let $\mathcal{B} = \mathcal{B}\big([-\|A\|, \|A\|]\big)$ denote the set of all bounded Borel functions on $[-\|A\|, \|A\|]$. There exists a unique map $\Phi : \mathcal{B} \to \mathcal{L}(\mathcal{H})$ such that
(a) Φ is an algebra $*$–homomorphism. That is

$$\begin{aligned} \Phi(\alpha f + \beta g) &= \alpha\Phi(f) + \beta\Phi(g) \qquad & \Phi(\bar{f}) &= \Phi(f)^* \\ \Phi(fg) &= \Phi(f)\Phi(g) & \Phi(1) &= \mathbb{1} \end{aligned}$$

for all $f, g \in \mathcal{B}$ and $\alpha, \beta \in \mathbb{C}$.

(b) $\big\|\Phi(f)\big\|_{\mathcal{L}(\mathcal{H})} \le \|f\|_{L^\infty}$ *for all $f \in \mathcal{B}$.*

(c) $\Phi(x) = A$

(d) If the sequence $\big\{f_n\big\}_{n\in\mathbb{N}} \subset \mathcal{B}$ converges pointwise to f and is uniformly bounded, then $\Phi(f) = \operatorname{s-lim}_{n\to\infty} \Phi(f_n)$.

Furthermore, Φ has the following properties.

(e) If, for some $\lambda \in \mathbb{R}$ and $\psi \in \mathcal{H}$, we have $A\psi = \lambda\psi$, then $\Phi(f)\psi = f(\lambda)\psi$, for all $f \in \mathcal{B}$.

(f) If $f \in \mathcal{B}$ is nonnegative, then $\Phi(f)$ is a nonnegative operator.

(g) If $A, B \in \mathcal{L}(\mathcal{H})$ commute, then so do $\Phi(f)$ and B, for all $f \in \mathcal{B}$. That is, if $AB = BA$, then $\Phi(f)B = B\Phi(f)$.

Remark 28 It is common to write $f(A)$ in place of $\Phi(f)$. In this notation, the above conclusions are

(a)

$$\begin{aligned} (\alpha f + \beta g)(A) &= \alpha f(A) + \beta g(A) & \bar{f}(A) &= f(A)^* \\ (fg)(A) &= f(A)g(A) & 1(A) &= \mathbb{1} \end{aligned}$$

(b) $\big\|f(A)\big\|_{\mathcal{L}(\mathcal{H})} \le \|f\|_{L^\infty}$

(c) $x(A) = A$

(d) If $f_n \to f$ pointwise and is uniformly bounded, then $f(A) = \operatorname*{s-lim}_{n\to\infty} f_n(A)$.

(e) If $A\psi = \lambda\psi$, then $f(A)\psi = f(\lambda)\psi$.

(f) If $f \ge 0$, then $f(A) \ge 0$.

(g) If $AB = BA$, then $f(A)B = Bf(A)$.

Example 29 Let A be a self–adjoint, compact operator and let $\big\{\varphi_n\big\}_{n\in\mathcal{I}}$ be a complete orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of A. Denote by λ_n the eigenvalue of A for the eigenvector φ_n. We just have to define, for each $f \in \mathcal{B}(\mathbb{R})$

$$f(A)\Big(\sum_{n\in\mathcal{I}} x_n\varphi_n\Big) = \sum_{n\in\mathcal{I}} f(\lambda_n)\, x_n\varphi_n$$

Definition 30 (Projection Valued Measure) Denote by $\mathcal{B}_{\mathbb{R}}$ the σ–algebra of Borel subsets of $\mathbb{R}$ and by $\mathcal{L}(\mathcal{H})$ the set of a bounded operators on $\mathcal{H}$. A projection valued measure is a map $E : \mathcal{B}_{\mathbb{R}} \to \mathcal{L}(\mathcal{H})$ that obeys the following conditions.

(i) For each $B \in \mathcal{B}_{\mathbb{R}}$, the operator $E(B)$ is an orthogonal projection on some closed subspace of $\mathcal{H}$. That is, $E(B)^2 = E(B)$ and $E(B) = E(B)^*$.

(ii) $E(\emptyset) = 0$ and $E(\mathbb{R}) = \mathbb{1}$

(iii) If $\{B_n\}_{n\in\mathbb{N}}$ is a countable family of disjoint Borel subsets of $\mathbb{R}$, then

$$E\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \operatorname*{s-lim}_{N\to\infty} \sum_{n=1}^{N} E(B_n)$$

A projection valued measure is said to be bounded if, in addition,

(iv) There is an $a > 0$ such that $E\big((-a, a)\big) = \mathbb{1}$.

A projection valued measure automatically also obeys

(v) $E(B_1 \cap B_2) = E(B_1)E(B_2)$ for all $B_1, B_2 \in \mathcal{B}_{\mathbb{R}}$.

(vi) $E(B_1)$ and $E(B_2)$ commute for all $B_1, B_2 \in \mathcal{B}_{\mathbb{R}}$.

Definition 31 **($\int f(\lambda)\ dE(\lambda)$)** Let $B \mapsto E(B)$ be a bounded projection valued measure, and $\varphi, \psi \in \mathcal{H}$. Observe that

- $B \mapsto \langle \varphi, E(B)\varphi\rangle$ is an ordinary finite Borel measure on $\mathbb{R}$.
- $B \mapsto \langle \varphi, E(B)\psi\rangle$ is an ordinary complex measure on $\mathbb{R}$. That is, there are (positive) measures μ_1, μ_2, ν_1 and ν_2 such that

$$\langle \varphi, E(B)\psi\rangle = \mu_1(B) - \mu_2(B) + i\nu_1(B) - i\nu_2(B)$$

By the polarization identity, we can take

$$\mu_1(B) = \tfrac{1}{4}\,\langle \varphi+\psi,\ E(B)\,(\varphi+\psi)\rangle \qquad \mu_2(B) = \tfrac{1}{4}\,\langle \varphi-\psi,\ E(B)\,(\varphi-\psi)\rangle$$
$$\nu_1(B) = \tfrac{1}{4}\,\langle \varphi+i\psi,\ E(B)\,(\varphi+i\psi)\rangle \quad \nu_2(B) = \tfrac{1}{4}\,\langle \varphi-i\psi,\ E(B)\,(\varphi-i\psi)\rangle$$

- Use $\mathcal{B}(\mathbb{R})$ to denote the set of bounded, Borel measurable functions on $\mathbb{R}$. If $f \in \mathcal{B}(\mathbb{R})$, then the map $(\varphi, \psi) \mapsto \int f(\lambda)\ d\,\langle \varphi, E(\lambda)\psi\rangle$ is well–defined, bounded and sesquilinear. So, by (a corollary to) the Riesz representation theorem, there is a unique $F \in \mathcal{L}(\mathcal{H})$ such that

$$\langle \varphi, F\psi\rangle = \int f(\lambda)\ d\,\langle \varphi, E(\lambda)\psi\rangle$$

for all $\varphi, \psi \in \mathcal{H}$.

We define

$$\int f(\lambda)\, dE(\lambda) = F$$

Example 32 Let $\mathcal{H} = L^2(M, \mu)$ for some measure space (M, μ) and let $a : M \to \mathbb{R}$ be bounded and measurable. Define, for each $B \in \mathcal{B}_{\mathbb{R}}$

$$E(B) = \text{multiplication by } \chi_{a^{-1}(B)}(m) = \chi_B\big(a(m)\big)$$

This is a bounded projection valued measure and

$$\int f(\lambda)\ d\,\langle \varphi, E(\lambda)\varphi\rangle = \int f\big(a(m)\big)\ |\varphi(m)|^2\ d\mu(m)$$
$$\int f(\lambda)\ d\,\langle \varphi, E(\lambda)\psi\rangle = \int f\big(a(m)\big)\ \overline{\varphi(m)}\psi(m)\ d\mu(m)$$

Theorem 33 (Spectral Theorem - Projection–valued Measure Version) *There is a 1–1 correspondence between bounded self–adjoint operators and bounded projection–valued measures $A \leftrightarrow E_A$ such that*

$$A = \int \lambda \, dE_A(\lambda)$$

Example 34 Let A be a self–adjoint, compact operator and let $\{\varphi_n\}_{n\in\mathcal{I}}$ be a complete orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of A. Denote by λ_n the eigenvalue of A for the eigenvector φ_n. Define, for each $m \in \mathcal{I}$, P_m to be the orthogonal projector onto the linear subspace of $\mathcal{H}$ consisting of all scalar multiples of φ_m. That is

$$P_m\Big(\sum_{n\in\mathcal{I}} x_n\varphi_n\Big) = x_m\varphi_m$$

Set, for each Borel subset B of $\mathbb{R}$

$$E_A(B) = \sum_{\substack{m\in\mathcal{I} \\ \text{with } \lambda_m\in B}} P_m$$

Then

$$A = \sum_{m\in\mathcal{I}} \lambda_m P_m = \sum_{m\in\mathcal{I}} \lambda_m E_A\big(\{\lambda_m\}\big) = \int \lambda \, dE_A(\lambda)$$

Corollary 35

$$\begin{aligned}
\lambda \in \sigma(A) &\iff E_A\big((\lambda-\varepsilon, \lambda+\varepsilon)\big) \neq 0 \qquad \textit{for all } \varepsilon > 0 \\
\lambda \in \sigma_p(A) &\iff E_A\big(\{\lambda\}\big) \neq 0 \\
E_A\big(\rho(A) \cap \mathbb{R}\big) &= 0 \\
\operatorname{range} E_A(\{0\}) &= \ker A
\end{aligned}$$

Corollary 36 *Let $-\infty < a < b < \infty$.*

(a) Then

$$\tfrac{1}{2}\big\{E\big([a,b]\big) + E\big((a,b)\big)\big\} = \operatorname*{s-lim}_{\varepsilon\to 0} \tfrac{1}{2\pi i} \int_a^b \big\{(A-\lambda-i\varepsilon)^{-1} - (A-\lambda+i\varepsilon)^{-1}\big\} \, d\lambda$$

(b) Let a and b be in the resolvent set of A. Then

$$E_A\big([a,b]\big) = E_A\big([a,b)\big) = E_A\big((a,b]\big) = E_A\big((a,b)\big) = \tfrac{1}{2\pi i} \int_{C_{a,b}} R_\zeta(A) \, d\zeta$$

for any simple closed curve $C_{a,b}$ in the complex plane with counterclockwise orientation that contains $\{\ \lambda + i0 \ \big|\ a \le \lambda \le b\ \}$ in its interior.

Definition 37 A Banach Algebra $\mathcal{A}$ is a set which

(a) is a complex algebra, i.e. a set equipped with three operations

$$(A,B)\in\mathcal{A}\times\mathcal{A}\mapsto A+B\in\mathcal{A} \quad (\alpha,A)\in\mathbb{C}\times\mathcal{A}\mapsto \alpha A\in\mathcal{A} \quad (A,B)\in\mathcal{A}\times\mathcal{A}\mapsto AB\in\mathcal{A}$$

called addition, scalar multiplication and multiplication, that obey the usual vector space axioms and

$$A(BC)=(AB)C \qquad (A+B)C=AB+BC \qquad A(B+C)=AB+AC$$
$$\alpha(BC)=(\alpha B)C=B(\alpha C)$$

for all $A,B,C\in\mathcal{A}$ and $\alpha\in\mathbb{C}$ (so multiplication is associative, but not necessarily commutative) and

(b) is normed, with the usual norm axioms, and also obeys $\|AB\|\le\|A\|\,\|B\|$ for all $A,B\in\mathcal{A}$ and

(c) and is complete.

If, in addition,

(d) $\mathcal{A}$ contains an identity element $\mathbb{1}$ that obeys $\|\mathbb{1}\|=1$ and $A\mathbb{1}=\mathbb{1}A=A$ for all $A\in\mathcal{A}$

then $\mathcal{A}$ is called a unital Banach algebra. Any Banach algebra can be easily extended to a unital Banach algebra $\mathcal{A}+\mathbb{1}$. (Define $\mathcal{A}_{\mathbb{1}}=\big\{\ (\alpha,A)\ \big|\ \alpha\in\mathbb{C},\ A\in\mathcal{A}\ \big\}$, the algebraic operations by thinking of (α,A) as $\alpha\mathbb{1}+A$, and $\|(\alpha,A)\|=|\alpha|+\|A\|_{\mathcal{A}}$.)

Definition 38 A C^*–algebra is a Banach algebra $\mathcal{A}$ together with a map $*:\mathcal{A}\to\mathcal{A}$ that obeys

$$(A+B)^*=A^*+B^* \qquad (\alpha A)^*=\bar\alpha A^* \qquad (AB)^*=B^*A^*$$
$$A^{**}=A \qquad\qquad \|A^*A\|=\|A\|^2$$

for all $A,B\in\mathcal{A}$ and $\alpha\in\mathbb{C}$. (This used to be called a B^*–algebra.)

Remark 39 The condition $\|A^*A\|=\|A\|^2$ implies that $\|A^*\|=\|A\|$ since

$$\|A\|^2=\|A^*A\|\le\|A^*\|\,\|A\| \implies \|A\|\le\|A^*\| \implies \|A^*\|\le\|A^{**}\|=\|A\|$$

Example 40

(a) Let $A=A^*\in\mathcal{L}(\mathcal{H})$. Then $\mathcal{A}=\overline{\big\{\text{polynomials in } A \text{ with complex coefficients}\big\}}$, with the overbar denoting norm closure, is a commutative C^*–algebra contained in $\mathcal{L}(\mathcal{H})$.

(b) Let $A_1,\ \cdots,\ A_n$ be self–adjoint, commuting, bounded operators on $\mathcal{H}$. Then $\mathcal{A}=\overline{\big\{\text{polynomials in } A_1,\ \cdots,\ A_n \text{ with complex coefficients}\big\}}$ is a commutative C^*–algebra in $\mathcal{L}(\mathcal{H})$.

(c) Any closed subalgebra of $\mathcal{L}(\mathcal{H})$ that is closed under the taking of adjoints is a C^*–algebra. Conversely, any C^*–algebra is isomorphic to a subalgebra of $\mathcal{L}(\mathcal{H})$.

(d) Let X be a Hausdorff space (that is, a set equipped with open sets such that distinct points have disjoint open neighbourhoods). Then $\mathcal{A} = C(X)$, the set of all continuous functions on X, with the supremum norm, is a commutative C^*–algebra .

Theorem 41 (Gelfand–Naimark)
Let $\mathcal{A}$ be a commutative C^*–algebra with identity. Then there is a compact Hausdorff space, X, (unique up to homeomorphism) such that $\mathcal{A}$ is $*$–isomorphic to $C(X)$. That is, there is a 1–1*, onto map $\Psi : \mathcal{A} \to C(X)$ such that*

$$\begin{array}{lll} \Psi(A+B) = \Psi(A) + \Psi(B) & \Psi(\alpha A) = \alpha\,\Psi(A) & \Psi(AB) = \Psi(A)\,\Psi(B) \\ \Psi(A^*) = \overline{\Psi(A)} & \big\|\Psi(A)\big\|_{C(X)} = \|A\|_{\mathcal{A}} & \end{array}$$

for all $A, B \in \mathcal{A}$ and $\alpha \in \mathbb{C}$.