8  The spectral theorem for self-adjoint operators

Theorem 8.1. (The spectral theorem for self-adjoint operators in finite dimension) For every self-adjoint \( n \times n \) matrix \( A \) there is an orthonormal basis of \( \mathbb{C}^n \) consisting of eigenvectors of \( A \); every eigenvalue is real.

Also in infinite-dimensional \( \mathcal{H} \), an eigenvector of \( A : \mathcal{D}(A) \to \mathcal{H} \) is a \( \psi \in \mathcal{D}(A) \setminus \{0\} \) such that
\[
A\psi = \lambda \psi
\]
for some \( \lambda \in \mathbb{C} \); then \( \lambda \) is called the eigenvalue of \( \psi \); a number \( \lambda \in \mathbb{C} \) is called an eigenvalue of \( A \) if there is a \( \psi \in \mathcal{D}(A) \setminus \{0\} \) such that (8.1) holds. If \( \mathcal{H} = L^2(\Omega) \) then eigenvectors are also called eigenfunctions. The set of eigenvectors with eigenvalue \( \lambda \), together with the zero vector, forms a subspace, called the eigenspace with eigenvalue \( \lambda \). If \( A \) is self-adjoint then its eigenvalues are real,
\[
\lambda \langle \psi | \psi \rangle = \langle \psi | A \psi \rangle = \langle A \psi | \psi \rangle = \lambda^* \langle \psi | \psi \rangle ,
\]
and eigenvectors corresponding to distinct eigenvalues are orthogonal:
\[
(\lambda_i - \lambda_j) \langle \psi_i | \psi_j \rangle = \langle A \psi_i | \psi_j \rangle - \langle \psi_i | A \psi_j \rangle = 0 .
\]

In infinite-dimensional \( \mathcal{H} \), operators do not necessarily have eigenvalues. For example, the free Hamiltonian has no eigenvalues, as it is unitarily equivalent to the multiplication operator \( M_{\omega} \), which has no eigenvalues:

Example 8.2. The eigenvalues of a multiplication operator \( M_{\omega}, V : \mathbb{R}^d \to \mathbb{C} \), are those values \( \lambda \in \mathbb{C} \) such that the set \( V^{-1}(\lambda) = \{ x \in \mathbb{R}^d : V(x) = \lambda \} \) has positive measure.

Indeed, if \( M_{\omega} \psi = \lambda \psi \) then \( V(x) \psi(x) = \lambda \psi(x) \) for all \( x \) except in a set of measure zero, and so \( V(x) = \lambda \) for all \( x \) with \( \psi(x) \neq 0 \) (except a null set). If \( \psi \neq 0 \) then \( \{ x : \psi(x) \neq 0 \} \) must have positive measure. Conversely, if \( V^{-1}(\lambda) \) has positive measure then any nonzero \( \psi \) that vanishes outside \( V^{-1}(\lambda) \) is an eigenfunction of \( M_{\omega} \) with eigenvalue \( \lambda \).

We thus need a notion of generalized eigenvalues; the set of generalized eigenvalues is called the spectrum.

Definition 8.3. The spectrum of the operator \( A : \mathcal{D}(A) \to \mathcal{H} \) is
\[
\sigma(A) = \{ z \in \mathbb{C} | (A - zI) : \mathcal{D}(A) \to \mathcal{H} \text{ is not bijective} \} .
\]

One breaks down the spectrum as follows:

- \( \sigma_p(A) = \{ z \in \mathbb{C} : A - zI \text{ is not injective} \} \)
  is called the point spectrum; it is the set of eigenvalues of \( A \).

- \( \sigma_c(A) = \{ z \in \mathbb{C} : A - zI \text{ is injective, not surjective, and has dense range} \} \)
  is called the continuous spectrum.
Theorem 8.5. Let $A$ be self-adjoint and densely defined in $\mathcal{H}$. Then $A$ has no residual spectrum, $\sigma(A) \subseteq \mathbb{R}$, and $\|A\| = \sup_{\lambda \in \sigma(A)} |\lambda|$ (called the spectral radius).


8.1 The spectral theorem in terms of multiplication operators

Theorem 8.6. Let $A$ be self-adjoint and densely defined in the separable Hilbert space $\mathcal{H}$. There is a measure space $(\Omega, \mathcal{A}, \mu)$ with finite measure $\mu$ and a measurable function $h : \Omega \to \mathbb{R}$ such that $A$ is unitarily equivalent to the multiplication operator $M_h$ on $L^2(\Omega, \mathcal{A}, \mu)$; i.e., there is a unitary $U : \mathcal{H} \to L^2(\Omega, \mathcal{A}, \mu)$ such that $\psi \in \mathcal{D}(A)$ iff $hU\psi \in L^2(\Omega, \mathcal{A}, \mu)$ and $UA\psi = hU\psi$ for all $\psi \in \mathcal{D}(A)$.


There is a lot of freedom in the choice of $\Omega, \mathcal{A}, \mu$, and $h$. Obviously, if $(\Omega, \mathcal{A}, \mu)$ is isomorphic to $(\Omega', \mathcal{A}', \mu')$ (i.e., if there is $\Phi : \Omega \to \Omega'$ that is measurable and bijective with measurable inverse and $\mu'(\Phi(\Delta)) = \mu(\Delta)$ for all $\Delta \in \mathcal{A}$), then $\Omega$ can be replaced with $\Omega'$ and $h$ with $h' = h \circ \Phi^{-1}$.

The space $\Omega$ can, in fact, be taken to be $\mathbb{R} \times \mathbb{N}$, or $\sigma(A) \times \mathbb{N}$, together with the function $h(x, n) = x$. The measure $\mu$, however, is usually not the Lebesgue measure.

For $\mathcal{H} = \mathbb{C}^n$, $\Omega$ can be taken to be a set of $n$ elements, $\mu = \#$, $h$ any mapping whose values are the eigenvalues of $A$ with appropriate multiplicity, and $\{U^{-1}e_i : i \in \Omega\}$, with $\{e_i\}$ the standard basis of $L^2(\Omega)$, an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of $A$. A multiplication operator then means a diagonal matrix. (In the representation $\Omega = \mathbb{R} \times \mathbb{N}$, and $\mathcal{H} = \mathbb{C}^n$, $\mu$ must be taken to be concentrated on $n$ points in $\Omega$, each with first component equal to an eigenvalue of $A$, and each eigenvalue occurring with appropriate multiplicity; $\mu$ can be taken to give equal weight to each of those $n$ points.)
8.2 The spectral theorem in terms of functional calculus

To define a functional calculus for an operator $A$ means to define operators $f(A)$ for all $f : \mathbb{R} \rightarrow \mathbb{C}$ or $f : \mathbb{C} \rightarrow \mathbb{C}$ or $f : \sigma(A) \rightarrow \mathbb{C}$ in some function space. If $f : \mathbb{C} \rightarrow \mathbb{C}$ is a polynomial, then it is obvious how to define $f(A)$ for $A \in \mathcal{B}(\mathcal{H})$; for self-adjoint unbounded $A$, $f(A)$ also makes sense but on a smaller domain:

$$\mathcal{D}(A^n) = \{ \psi \in \mathcal{H} : \psi \in \mathcal{D}(A), A\psi \in \mathcal{D}(A), A^2\psi \in \mathcal{D}(A), \ldots, A^{n-1}\psi \in \mathcal{D}(A) \} \quad (8.6)$$

It follows from the spectral theorem 8.6 that

$$\mathcal{D}(A^n) = \{ \psi \in \mathcal{H} : h^nU\psi \in L^2(\Omega, \mathfrak{A}, \mu) \} . \quad (8.7)$$

For $A \in \mathcal{B}(\mathcal{H})$, one can also easily define $f(A)$ if $f : \mathbb{C} \rightarrow \mathbb{C}$ is an entire holomorphic function, i.e., given by a power series of infinite radius of convergence: If

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \text{ then } f(A) = \sum_{n=0}^{\infty} c_n A^n \quad (8.8)$$

is norm convergent.

In $\mathcal{H} = \mathbb{C}^n$, the functional calculus for self-adjoint $A$ can be defined for arbitrary functions $f$ by diagonalization: If $A = U^{-1}\text{diag}(\lambda_1, \ldots, \lambda_n)U$ then set

$$f(A) = U^{-1}\text{diag}(f(\lambda_1), \ldots, f(\lambda_n))U . \quad (8.9)$$

This functional calculus obviously extends the one for polynomials and power series.

So, different approaches allow us to define $f(A)$ for different types of operators $A$ and different classes of functions $f$. Let $\mathcal{L}^\infty(\mathbb{R})$ be the space of bounded measurable functions; note $L^\infty = \mathcal{L}^\infty / \sim$, where $\sim$ means equality almost everywhere.

Theorem 8.7. (Spectral theorem in terms of functional calculus) Let $A$ be self-adjoint and densely defined in the separable Hilbert space $\mathcal{H}$. Then there is a unique mapping $\mathcal{L}^\infty(\mathbb{R}) \rightarrow \mathcal{B}(\mathcal{H})$, denoted $f \mapsto f(A)$ and called the functional calculus of $A$, such that

(i) $f \mapsto f(A)$ is a homomorphism of algebras, i.e., linear and multiplicative:

$$ (f + \lambda g)(A) = f(A) + \lambda g(A) , \quad (fg)(A) = f(A)g(A) \quad \forall f, g \in E . \quad (8.10) $$

(ii) $f^*(A) = f(A)^*$

(iii) $\|f(A)\| \leq \|f\|_\infty$ (where $\| \cdot \|_\infty$ is the supremum, not the essential supremum)

(iv) If a sequence $f_n \in \mathcal{L}^\infty(\mathbb{R})$ converges pointwise to $f$ and $\|f_n(x)\| \leq |x|$ for all $x$ and $n$ then $\lim_{n \rightarrow \infty} f_n(A)\psi = A\psi$ for every $\psi \in \mathcal{D}(A)$.

(v) If a sequence $f_n \in \mathcal{L}^\infty(\mathbb{R})$ converges pointwise to $f \in \mathcal{L}^\infty(\mathbb{R})$ and $\sup_n \|f_n\|_\infty < \infty$ then $\text{s-lim}_{n \rightarrow \infty} f_n(A) = f(A)$. 

44
In addition:

(vi) If \( A\psi = \lambda\psi \) then \( f(A)\psi = f(\lambda)\psi \).

(vii) If \( f \) vanishes on the spectrum of \( A \) then \( f(A) = 0 \).

Proof. Uniqueness: see Reed and Simon vol. I, Theorems VII.1, VII.2, VIII.5.

Existence: follows from the first version of the spectral theorem 8.6 by setting

\[
f(A) = U^{-1}M_{f\circ h}U.
\]  

(8.11)

Since \( f \circ h \) is bounded, \( M_{f\circ h} \in \mathcal{B}(\mathcal{H}) \). Now the properties can be verified.

\( \square \)

**Corollary 8.8.** Let \( \mathcal{H} \) be separable, \( H \) self-adjoint and densely defined, and \( U_t = e^{-ixt/\hbar} \) as defined by the functional calculus. Then \( U_t \) is a strongly continuous unitary group with generator \( H \).

**Proof.** The group property and \( U_t^{-1} = U_{-t} = U_t^* \) follow from (i) and (ii). Furthermore, \( e^{-ixt} \to 1 \) pointwise as \( t \to 0 \) and \( \|e^{-ixt}\|_\infty = 1 \); by (v), \( \lim_{t \to 0} e^{-ixt/\hbar} = I \); thus, \( U_t \) is a strongly continuous unitary group.

To see that \( H \) is the generator, let \( U : \mathcal{H} \to L^2(\Omega, \mathfrak{A}, \mu) \) be as in the spectral theorem 8.6 for \( A = H \). Then \( U_t = U^{-1}e^{-ixt/\hbar}U \) and

\[
\frac{id}{dt}U_t\psi = i\hbar U^{-1}\frac{d}{dt}e^{-ixt/\hbar}U\psi = U^{-1}he^{-ixt/\hbar}U\psi \in \mathcal{H} \Leftrightarrow hU\psi \in L^2(\Omega)
\]  

\( \Leftrightarrow H\psi \in \mathcal{H} \Leftrightarrow \psi \in \mathcal{D}(H) \).  

(8.12)

(8.13)

Thus, \( H \) is a generator of \( U_t \).

\( \square \)

**8.3 The spectral theorem in terms of PVMs**

Another formulation of the spectral theorem in \( \mathbb{C}^n \) says that any self-adjoint \( A \) can be written as

\[
A = \sum_{j=1}^{m} \lambda_j P_j,
\]  

(8.14)

where the \( \lambda_j \) are the eigenvalues of \( A \), \( m \) is the number of eigenvalues, and \( P_j \) is the projection to the eigenspace of \( \lambda_j \). In fact, this is the unique decomposition of \( A \) as a real-linear combination of projections that are mutually orthogonal. This formulation can be transferred to the infinite-dimensional case.

**Definition 8.9.** A *projection-valued measure* (PVM) \( P \) on a measurable space \( (\Omega, \mathfrak{A}) \) acting on \( \mathcal{H} \) is a mapping \( \mathfrak{A} \to \mathcal{B}(\mathcal{H}) \) such that

- for every \( \Delta \in \mathfrak{A} \), \( P(\Delta) \) is a projection,
- \( P(\Omega) = I \),

45
Corollary 8.10. (i) $P(\emptyset) = 0$.

(ii) $P$ is also finitely additive, e.g., $P(\Delta_1 \cup \Delta_2) = P(\Delta_1) + P(\Delta_2)$ if $\Delta_1 \cap \Delta_2 = \emptyset$.

(iii) If $\Delta_1 \cap \Delta_2 = \emptyset$ then $P(\Delta_1)$ and $P(\Delta_2)$ correspond to orthogonal subspaces, $P(\Delta_1)P(\Delta_2) = 0$.

(iv) $P(\Delta_1 \cap \Delta_2) = P(\Delta_1)P(\Delta_2)$

(v) The series (8.15) also converges strongly, i.e., $\sum_n P(\Delta_n)\psi$ converges for every $\psi \in \mathcal{H}$.

Proof. (i) Set $\Delta_1 = \Delta_2 = \ldots = \emptyset$ and obtain $P(\emptyset) = P(\emptyset) + P(\emptyset) + \ldots$, which can converge weakly only if $\langle \psi | P(\emptyset)\psi \rangle = 0$ for every $\psi \in \mathcal{H}$, which implies $P(\emptyset) = 0$.

(ii) Set $\Delta_3 = \Delta_4 = \ldots = \emptyset$.

(iii) If the sum of two projections is again a projection, they must be mutually orthogonal: Recall that $P^2 = P$ for any projection $P$. If $P_1 + P_2$ is a projection then

$$P_1 + P_2 = (P_1 + P_2)^2 = P_1^2 + P_1P_2 + P_2P_1 + P_2^2 = P_1 + P_2 + P_1P_2 + P_2P_1$$

so $P_1P_2 + P_2P_1 = 0$. For any $\psi_1$ in the range of $P_1$, it follows that $P_1P_2\psi = -P_2\psi$, but since $P_1\phi = -\phi$ implies $\phi = 0$, we have that $P_2\psi_1 = 0$. In the same way, we see $P_1\psi_2 = 0$ for any $\psi_2$ in the range of $P_2$, so the two subspaces are orthogonal, and $P_1P_2 = 0 = P_2P_1$.

(iv) Set $A = \Delta_1 \cap \Delta_2$, $B = \Delta_1 \setminus A$, $C = \Delta_2 \setminus A$. Then

$$P(\Delta_1)P(\Delta_2) = P(A \cup B)P(A \cup C) = (P(A) + P(B))(P(A) + P(C))$$

$$= P(A)^2 + P(B)P(A) + P(A)P(C) + P(B)P(C) = P(A).$$

(v) Set $\Delta = \bigcup_{n=1}^{\infty} \Delta_n$ and $\tilde{\Delta}_N = \bigcup_{n=N+1}^{\infty} \Delta_n$.

$$\left\| \sum_{n=1}^{N} P(\Delta_n)\psi - P(\Delta)\psi \right\|^2 = \left\| P\left( \bigcup_{n=1}^{N} \Delta_n \right)\psi - P(\Delta)\psi \right\|^2 = \left\| P(\tilde{\Delta}_N)\psi \right\|^2$$

$$= \langle \psi | P(\tilde{\Delta}_N)^2\psi \rangle = \langle \psi | P(\tilde{\Delta}_N)\psi \rangle = \sum_{n=1}^{N} \langle \psi | P(\Delta_n)\psi \rangle - \langle \psi | P(\Delta)\psi \rangle \xrightarrow{N \to \infty} 0.$$
Example 8.11. For any measure space \((\Omega, \mathcal{A}, \mu)\), the natural PVM on \(\Omega\) acting on \(L^2(\Omega, \mathcal{A}, \mu)\) is \(P(\Delta) = M_\Delta\), i.e.,

\[
(P(\Delta)\psi)(\omega) = \begin{cases} 
\psi(\omega) & \text{if } \omega \in \Delta \\
0 & \text{if } \omega \notin \Delta.
\end{cases}
\quad (8.18)
\]

Example 8.12. Suppose that \(\Omega\) is a finite or countable set. Then \(P(\Delta)\) can be expressed by singletons:

\[
P(\Delta) = \sum_{j \in \Delta} P\{j\}. \quad (8.19)
\]

Thus, the PVM \(P(\cdot)\) is determined by the \(P_j = P(\{j\})\). Conversely, if \(\{P_j : j \in \Omega\}\) is a family of mutually orthogonal projections with \(\sum_{j \in \Omega} P_j = I\) then (8.19) defines a PVM. As a consequence, the diagonalization of a matrix \(A : \mathbb{C}^n \to \mathbb{C}^n\) as in (8.14) defines a PVM on \(\sigma(A)\) acting on \(\mathbb{C}^n\); it can also be regarded as a PVM on \(\mathbb{R}\); \(P(\Delta)\) is the projection to the subspace spanned by the eigenspaces of all eigenvalues contained in \(\Delta\).

From the functional calculus it follows that with a self-adjoint operator \(A\) on a separable \(\mathcal{H}\) there is associated a spectral PVM \(P(\cdot)\) on \(\mathbb{R}\) acting on \(\mathcal{H}\) by

\[
P(\Delta) = 1_\Delta(A).
\quad (8.20)
\]

**Theorem 8.13.** (Spectral theorem in terms of PVMs) Let \(A\) be self-adjoint and densely defined in a separable \(\mathcal{H}\). There is a unique PVM \(P(\cdot)\) on \(\mathbb{R}\) acting on \(\mathcal{H}\) such that

\[
A = \int_{\mathbb{R}} x P(dx).
\quad (8.21)
\]

\(P\) is the spectral PVM of \(A\). Conversely, the right hand side of (8.21) is always a self-adjoint and densely defined operator.

**Definition 8.14.** For a given PVM \(P\), we define the operator

\[
B_f = \int_{\Omega} f(\omega) P(d\omega)
\quad (8.22)
\]

for any measurable \(f : \Omega \to \mathbb{C}\) by “weak integration”: For any \(\psi, \phi \in \mathcal{H}\),

\[
\mu_{\psi, \phi}(\Delta) = \langle \psi | P(\Delta) \phi \rangle
\quad (8.23)
\]

defines a complex measure \(\mu_{\psi, \phi}\) on \(\Omega\); we define \(B_f\) by the property

\[
\langle \psi | B_f \phi \rangle = \int_{\Omega} f(\omega) \mu_{\psi, \phi}(d\omega) \quad \forall \psi, \phi \in \mathcal{D}(B_f)
\quad (8.24)
\]

on the domain

\[
\mathcal{D}(B_f) = \left\{ \phi \in \mathcal{H} : \int_{\Omega} |f(\omega)|^2 \mu_{\phi, \phi}(d\omega) < \infty \right\}.
\quad (8.25)
\]
**Proposition 8.15.** \( \mathcal{D}(B_f) \) is a dense subspace, and (8.24) defines a unique operator \( B_f \) on \( \mathcal{D}(B_f) \).

**Proof.** We first show that \( \mathcal{D}(B_f) \) is a subspace. Suppose \( \psi, \phi \in \mathcal{D}(B_f) \); \( \lambda \phi \in \mathcal{D}(B_f) \) is clear; we need to show that \( \psi + \phi \in \mathcal{D}(B_f) \). Let \( f_n(\omega) = \sum_{j=1}^{n} \lambda_j \Delta \omega \) be a sequence of simple functions converging pointwise monotonically to \( f(\omega) \).

\[
\int |f|^2 \mu_\psi(d\omega) = \lim_{n \to \infty} \int |f_n|^2 \mu_\psi(d\omega) = \lim_{n \to \infty} \sum_{j=1}^{n} \lambda_j^2 \langle \psi, P(\Delta_j) \rangle \langle \phi, P(\Delta_j) \rangle
\]

\[
\leq \lim_{n \to \infty} \sum_{j=1}^{n} \lambda_j^2 \left( \langle \psi, P(\Delta_j) \rangle \right)^{\frac{1}{2}} \left( \langle \phi, P(\Delta_j) \rangle \right)^{\frac{1}{2}}
\]

\[
= \left( \int |f|^2 \mu_\psi(d\omega) \right)^{\frac{1}{2}} \left( \int |f|^2 \mu_\phi(d\omega) \right)^{\frac{1}{2}} < \infty.
\]

(8.26)

since, by the Cauchy–Schwarz inequality in \( \mathbb{R}^n \),

\[
\lim_{n \to \infty} \sum_{j=1}^{n} \lambda_j^2 \langle \psi, P(\Delta_j) \rangle \langle \phi, P(\Delta_j) \rangle \leq \left( \sum_{j=1}^{n} \lambda_j^2 \langle \psi, P(\Delta_j) \rangle \right)^{\frac{1}{2}} \left( \sum_{j=1}^{n} \lambda_j^2 \langle \phi, P(\Delta_j) \rangle \right)^{\frac{1}{2}}
\]

\[
= \left( \int |f|^2 \mu_\psi(d\omega) \right)^{\frac{1}{2}} \left( \int |f|^2 \mu_\phi(d\omega) \right)^{\frac{1}{2}} < \infty.
\]

(8.27)

Thus, \( \mathcal{D}(B_f) \) is a subspace. We now show it is dense: For any \( \psi \in \mathcal{H} \) and \( n \in \mathbb{N} \), let

\[
\tilde{\Delta}_n = \{ \omega \in \Omega : |f(\omega)|^2 < n \}
\]

and \( \psi_n = P(\tilde{\Delta}_n) \psi \). Then \( \psi_n \in \mathcal{D}(B_f) \) because \( \mu_{\psi_n} \) is concentrated on \( \tilde{\Delta}_n \) because

\[
\mu_{\psi_n}(\Delta) = \langle \psi, P(\tilde{\Delta}_n) \rangle \langle \psi, P(\Delta) \rangle = \langle \psi, P(\Delta \cap \tilde{\Delta}_n) \rangle
\]

by Corollary 8.10(iv). Furthermore, \( \psi_n \to \psi \) by Corollary 8.10(v) for \( \Delta_n = \{ n-1 < |f|^2 < n \} \).

To see that \( \int f \mu_\psi(d\omega) \) is well defined, we use polarization,

\[
\mu_\psi = \frac{1}{4} \left( \mu_{\psi+\phi,\psi-\phi} - \mu_{\psi-\phi,\psi+\phi} + i \mu_{\psi-i\phi,\psi-i\phi} - i \mu_{\psi+i\phi,\psi+i\phi} \right),
\]

(8.30)
and note that for any $\chi \in \{ \psi + \phi, \psi - \phi, \psi - i\phi, \psi + i\phi \}$ we have that $\chi \in \mathcal{D}(B_f)$ and thus

$$
\int |f| |\mu_{\chi}(d\omega)| \leq \int \frac{1}{\Delta_1} |\mu_{\chi}(d\omega)| + \int |f|^2 |\mu_{\chi}(d\omega)| < \infty.
$$

(8.31)

It now follows that the rhs of (8.24) is a sesquilinear form $S(\psi, \phi)$. We now show that for every $\phi \in \mathcal{D}(B_f)$ there is $C_\phi > 0$ such that, for all $\psi \in \mathcal{D}(B_f)$,

$$
|S(\psi, \phi)| \leq C_\phi \|\psi\|.
$$

(8.32)

It then follows from the Riesz representation theorem, by regarding $\phi$ as fixed and $\psi$ as a variable, that $S(\psi, \phi) = \langle \psi | \chi_{\phi} \rangle$ for some unique $\chi_{\phi} \in \mathcal{H}$; since $\chi_{\phi}$ depends linearly on $\phi$, it defines an operator $B_f \phi = \chi_{\phi}$. To see (8.32), choose simple functions $f_n = \sum_{j=1}^n \lambda_{jn} \Delta_j \psi$ such that $f_n \to f$ pointwise and $|f_n| \leq |f|$. Then

$$
|S(\psi, \phi)| = \left| \int f(\omega) \mu_{\psi,\phi}(d\omega) \right|
$$

(8.33)

$$
= \lim_{n \to \infty} \left| \int f_n(\omega) \mu_{\psi,\phi}(d\omega) \right|
$$

(8.34)

$$
= \lim_{n \to \infty} \left| \sum_{j=1}^n \lambda_{jn} \frac{\mu_{\psi,\phi}(\Delta_j \psi)}{\langle \Delta_j \psi | \psi \rangle} \right|
$$

(8.35)

$$
\leq \lim_{n \to \infty} \sum_{j=1}^n |\lambda_{jn}| \|P(\Delta_j \psi)\| \|P(\Delta_j \phi)\|
$$

(8.36)

$$
\leq \lim_{n \to \infty} \left( \sum_{j=1}^n \|P(\Delta_j \psi)\|^2 \right)^{1/2} \left( \sum_{j=1}^n |\lambda_{jn}|^2 \|P(\Delta_j \phi)\|^2 \right)^{1/2}
$$

(8.37)

$$
= \lim_{n \to \infty} \|\psi\| \left( \int |f_n|^2 \mu_{\phi,\phi}(d\omega) \right)^{1/2}.
$$

(8.38)

The functional calculus can be recovered from the spectral PVM of $A$ by defining

$$
f(A) = \int_{\mathbb{R}} f(x) P(dx).
$$

(8.39)
By a diagonalization of $A$, we mean either the unitary equivalence $U : \mathcal{H} \to L^2(\Omega, \mathfrak{A}, \mu)$ that carries $A$ into a multiplication operator, or the PVM such that $A = \int x P(dx)$.

We say that the self-adjoint operators $A_1, \ldots, A_n$ are simultaneously diagonalizable if there is one $U : \mathcal{H} \to L^2(\Omega, \mathfrak{A}, \mu)$ that will carry each $A_i$ into a multiplication operator $M_{h_i}$; equivalently, there is one PVM $P(\cdot)$ on $\mathbb{R}^n$ such that

$$A_i = \int x_i P(dx) \quad (8.40)$$

for each $i$. We report that this is the case iff the $A_i$ commute pairwise. For bounded operators, it is clear what that means: $A_i A_j = A_j A_i$. For unbounded operators, the meaning is less clear because $A_j \psi$ may not lie in the domain of $A_i$. That is why one says that two unbounded operators $A_1, A_2$ commute iff $P_1(\Delta_1) P_2(\Delta_2) = P_2(\Delta_2) P_1(\Delta_1)$ for all measurable sets $\Delta_1, \Delta_2 \subseteq \mathbb{R}$, with $P_i$ the spectral PVM of $A_i$.

Can non-self-adjoint operators be diagonalized? Not necessarily. Let us consider only bounded operators. Note first that every $A \in \mathcal{B}(\mathcal{H})$ can be written in a unique way as

$$A = B + iC \quad (8.41)$$

with self-adjoint $B, C \in \mathcal{B}(\mathcal{H})$; indeed, $B = \frac{1}{2}(A + A^*)$ and $C = \frac{i}{2}(A - A^*)$. $A$ can be diagonalized iff $B$ and $C$ can be simultaneously diagonalized, which occurs iff $B$ and $C$ commute. This occurs iff $A$ and $A^*$ commute; such operators are called normal. For example, every unitary is normal because $U^* U = I = U U^*$.

A PVM can provide a notion of generalized orthonormal basis that we hinted at before. An ordinary orthonormal basis $\{\phi_n : n \in \mathbb{N}\}$ would correspond to the PVM $P$ on $\mathbb{N}$ given by

$$P(\Delta) = \sum_{n \in \Delta} P_{\phi_n}, \quad (8.42)$$

which is the projection to $\text{span}\{\phi_n : n \in \Delta\}$. In Section 3.1 we defined a generalized ONB as a unitary $U : \mathcal{H} \to L^2(\Omega, \mathfrak{A}, \mu)$; define the PVM $P$ on $\Omega$ by $P(\Delta) = U M_{1_\Delta} U^*$.

I will from now on take “self-adjoint” to include “densely defined.”