CHAPTER 2

Hilbert spaces and operators

1. Hilbert space

We have shown that $L^p(X,\mu)$ is a Banach space – a complete normed space. I shall next discuss the class of Hilbert spaces, a special class of Banach spaces, of which $L^2(X,\mu)$ is a standard example, in which the norm arises from an inner product, just as it does in Euclidean space.

An inner product on a vector space $V$ over $\mathbb{C}$ (one can do the real case too, not much changes) is a sesquilinear form

$$V \times V \to \mathbb{C}$$

written $(u,v)$, if $u,v \in V$. The ‘sesqui-’ part is just linearity in the first variable

$$\text{(1.1)} \quad (a_1 u_1 + a_2 u_2, v) = a_1 (u_1, v) + a_2 (u_2, v),$$

anti-linearly in the second

$$\text{(1.2)} \quad (u, a_1 v_1 + a_2 v_2) = \overline{a_1} (u, v_1) + \overline{a_2} (u, v_2)$$

and the conjugacy condition

$$\text{(1.3)} \quad (u, v) = \overline{(v, u)}.$$

Notice that (1.2) follows from (1.1) and (1.3). If we assume in addition the positivity condition\(^1\)

$$\text{(1.4)} \quad (u, u) \geq 0, \quad (u, u) = 0 \Rightarrow u = 0,$$

then

$$\text{(1.5)} \quad \|u\| = (u, u)^{1/2}$$

is a norm on $V$, as we shall see.

Suppose that $u,v \in V$ have $\|u\| = \|v\| = 1$. Then $(u,v) = e^{i\theta}|(u,v)|$ for some $\theta \in \mathbb{R}$. By choice of $\theta$, $e^{-i\theta}(u,v) = |(u,v)|$ is

\(^{1}\)Notice that $(u,u)$ is real by (1.3).
real, so expanding out using linearity for \( s \in \mathbb{R} \),
\[
0 \leq (e^{-i\theta}u - sv, e^{-i\theta}u - sv)
= \|u\|^2 - 2s \text{Re } e^{-i\theta}(u, v) + s^2\|v\|^2 = 1 - 2s|(u, v)| + s^2.
\]
The minimum of this occurs when \( s = |(u, v)| \) and this is negative unless \(|(u, v)| \leq 1 \). Using linearity, and checking the trivial cases \( u = v = 0 \) shows that
\[
|(u, v)| \leq \|u\|\|v\|, \forall u, v \in V.
\]
This is called Schwarz’\(^2\) inequality.

Using Schwarz’ inequality
\[
\|u + v\|^2 = \|u\|^2 + (u, v) + (v, u) + \|v\|^2
\leq (\|u\| + \|v\||)^2
\implies \|u + v\| \leq \|u\| + \|v\| \forall u, v \in V
\]
which is the triangle inequality.

**Definition 1.1.** A Hilbert space is a vector space \( V \) with an inner product satisfying (1.1) - (1.4) which is complete as a normed space (i.e., is a Banach space).

Thus we have already shown \( L^2(X, \mu) \) to be a Hilbert space for any positive measure \( \mu \). The inner product is
\[
(f, g) = \int_X f\overline{g} \, d\mu,
\]
since then (1.3) gives \( \|f\|_2 \).

Another important identity valid in any inner product spaces is the parallelogram law:
\[
\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.
\]
This can be used to prove the basic ‘existence theorem’ in Hilbert space theory.

**Lemma 1.2.** Let \( C \subset H \), in a Hilbert space, be closed and convex (i.e., \( su + (1 - s)v \in C \) if \( u, v \in C \) and \( 0 < s < 1 \)). Then \( C \) contains a unique element of smallest norm.

**Proof.** We can certainly choose a sequence \( u_n \in C \) such that \( \|u_n\| \to \delta = \inf \{\|v\| \mid v \in C\} \).

\(^2\)No ‘t’ in this Schwarz.
By the parallelogram law,
\[ \|u_n - u_m\|^2 = 2\|u_n\|^2 + 2\|u_m\|^2 - \|u_n + u_m\|^2 \]
\[ \leq 2(\|u_n\|^2 + \|u_m\|^2) - 4\delta^2 \]
where we use the fact that \((u_n + u_m)/2 \in C\) so must have norm at least \(\delta\). Thus \(\{u_n\}\) is a Cauchy sequence, hence convergent by the assumed completeness of \(H\). Thus \(\lim u_n = u \in C\) (since it is assumed closed) and by the triangle inequality
\[ |\|u_n\| - \|u\|| \leq \|u_n - u\| \to 0 \]
So \(\|u\| = \delta\). Uniqueness of \(u\) follows again from the parallelogram law which shows that if \(\|u'\| = \delta\) then
\[ \|u - u'\| \leq 2\delta^2 - 4\|u + u'\|/2 \leq 0 . \]

\[ \square \]

The fundamental fact about a Hilbert space is that each element \(v \in H\) defines a continuous linear functional by
\[ H \ni u \mapsto (u, v) \in \mathbb{C} \]
and conversely every continuous linear functional arises this way. This is also called the Riesz representation theorem.

**Proposition 1.3.** If \(L : H \to \mathbb{C}\) is a continuous linear functional on a Hilbert space then this is a unique element \(v \in H\) such that
\[ (1.9) \quad Lu = (u, v) \ \forall \ u \in H , \]

**Proof.** Consider the linear space
\[ M = \{u \in H ; Lu = 0\} \]
the null space of \(L\), a continuous linear functional on \(H\). By the assumed continuity, \(M\) is closed. We can suppose that \(L\) is not identically zero (since then \(v = 0\) in (1.9)). Thus there exists \(w \notin M\). Consider
\[ w + M = \{v \in H ; v = w + u , u \in M\} . \]
This is a closed convex subset of \(H\). Applying Lemma 1.2 it has a unique smallest element, \(v \in w + M\). Since \(v\) minimizes the norm on \(w + M\),
\[ \|v + su\|^2 = \|v\|^2 + 2\text{Re}(sv, v) + \|s\|^2\|u\|^2 \]
is stationary at \(s = 0\). Thus \(\text{Re}(u, v) = 0 \ \forall \ u \in M\), and the same argument with \(s\) replaced by \(i\) shows that \((v, u) = 0 \ \forall \ u \in M\).

Now \(v \in w + M\), so \(Lv = Lw \neq 0\). Consider the element \(w' = w/Lw \in H\). Since \(Lw' = 1\), for any \(u \in H\)
\[ L(u - (Lu)w') = Lu - Lu = 0 . \]
It follows that \( u - (Lu)w' \in M \) so if \( w'' = w'/\|w'\|^2 \)

\[
(u, w'') = ((Lu)w', w'') = Lu \frac{(w', w')}{\|w'\|^2} = Lu.
\]

The uniqueness of \( v \) follows from the positivity of the norm. \( \square \)

**Corollary 1.4.** For any positive measure \( \mu \), any continuous linear functional

\[ L : L^2(X, \mu) \to \mathbb{C} \]

is of the form

\[ Lf = \int_X f\overline{g} \, d\mu, \quad g \in L^2(X, \mu). \]

Notice the apparent power of ‘abstract reasoning’ here! Although we seem to have constructed \( g \) out of nowhere, its existence follows from the completeness of \( L^2(X, \mu) \), but it is very convenient to express the argument abstractly for a general Hilbert space.

**2. Spectral theorem**

For a bounded operator \( T \) on a Hilbert space we define the spectrum as the set

\[
\text{spec}(T) = \{ z \in \mathbb{C}; T - z \text{ Id is not invertible} \}. \tag{2.1}
\]

**Proposition 2.1.** For any bounded linear operator on a Hilbert space \( \text{spec}(T) \subset \mathbb{C} \) is a compact subset of \( \{ |z| \leq \|T\| \} \).

**Proof.** We show that the set \( \mathbb{C} \setminus \text{spec}(T) \) (generally called the resolvent set of \( T \)) is open and contains the complement of a sufficiently large ball. This is based on the convergence of the Neumann series. Namely if \( T \) is bounded and \( \|T\| < 1 \) then

\[
(Id - T)^{-1} = \sum_{j=0}^{\infty} T^j \tag{2.2}
\]

converges to a bounded operator which is a two-sided inverse of \( \text{Id} - T \). Indeed, \( \|T^j\| \leq \|T\|^j \) so the series is convergent and composing with \( \text{Id} - T \) on either side gives a telescoping series reducing to the identity.

Applying this result, we first see that

\[
(T - z) = -z(\text{Id} - T/z) \tag{2.3}
\]

is invertible if \( |z| > \|T\| \). Similarly, if \( (T - z_0)^{-1} \) exists for some \( z_0 \in \mathbb{C} \) then

\[
(T - z) = (T - z_0) - (z - z_0) = (T - z_0)^{-1}(\text{Id} - (z - z_0)(T - z_0)^{-1}) \tag{2.4}
\]

exists for \( |z - z_0|\|(T - z_0)^{-1}\| < 1. \) \( \square \)
In general it is rather difficult to precisely locate \( \text{spec}(T) \).

However for a bounded self-adjoint operator it is easier. One sign of this is the the norm of the operator has an alternative, simple, characterization. Namely

(2.5) \[ \text{if } A^* = A \text{ then } \sup_{\|\phi\| = 1} \langle A\phi, \phi \rangle = \|A\|. \]

If \( a \) is this supermum, then clearly \( a \leq \|A\| \). To see the converse, choose any \( \phi, \psi \in H \) with norm 1 and then replace \( \psi \) by \( e^{i\theta} \psi \) with \( \theta \) chosen so that \( \langle A\phi, \psi \rangle \) is real. Then use the polarization identity to write

(2.6) \[ 4\langle A\phi, \psi \rangle = \langle A(\phi + \psi), (\phi + \psi) \rangle - \langle A(\phi - \psi), (\phi - \psi) \rangle \\
+ i\langle A(\phi + i\psi), (\phi + i\psi) \rangle - i\langle A(\phi - i\psi), (\phi - i\psi) \rangle. \]

Now, by the assumed reality we may drop the last two terms and see that

(2.7) \[ 4|\langle A\phi, \psi \rangle| \leq a(\|\phi + \psi\|^2 + \|\phi - \psi\|^2) = 2a(\|\phi\|^2 + \|\psi\|^2) = 4a. \]

Thus indeed \( \|A\| = \sup_{\|\phi\| = \|\psi\| = 1} |\langle A\phi, \psi \rangle| = a. \)

We can always subtract a real constant from \( A \) so that \( A' = A - t \) satisfies

(2.8) \[ -\inf_{\|\phi\| = 1} \langle A'\phi, \phi \rangle = \sup_{\|\phi\| = 1} \langle A'\phi, \phi \rangle = \|A'\|. \]

Then, it follows that \( A' \pm \|A'\| \) is not invertible. Indeed, there exists a sequence \( \phi_n \), with \( \|\phi_n\| = 1 \) such that \( \langle (A' - \|A'\|)\phi_n, \phi_n \rangle \to 0 \). Thus

(2.9) \[ \| (A' - \|A'\|)\phi_n \|^2 = -2\langle A'\phi_n, \phi_n \rangle + \|A'\phi_n\|^2 + \|A'\|^2 \leq -2\langle A'\phi_n, \phi_n \rangle + 2\|A'\|^2 \to 0. \]

This shows that \( A' - \|A'\| \) cannot be invertible and the same argument works for \( A' + \|A'\| \). For the original operator \( A \) if we set

(2.10) \[ m = \inf_{\|\phi\| = 1} \langle A\phi, \phi \rangle \quad M = \sup_{\|\phi\| = 1} \langle A\phi, \phi \rangle \]

then we conclude that neither \( A - m \text{ Id} \) nor \( A - M \text{ Id} \) is invertible and \( \|A\| = \max(-m, M) \).

**Proposition 2.2.** If \( A \) is a bounded self-adjoint operator then, with \( m \) and \( M \) defined by (2.10),

(2.11) \[ \{m\} \cup \{M\} \subset \text{spec}(A) \subset [m, M]. \]

**Proof.** We have already shown the first part, that \( m \) and \( M \) are in the spectrum so it remains to show that \( A - z \) is invertible for all \( z \in \mathbb{C} \setminus [m, M] \).

Using the self-adjointness

(2.12) \[ \text{Im} \langle (A - z)\phi, \phi \rangle = -\text{Im} z \|\phi\|^2. \]
This implies that $A - z$ is invertible if $z \in \mathbb{C} \setminus \mathbb{R}$. First it shows that $(A - z)\phi = 0$ implies $\phi = 0$, so $A - z$ is injective. Secondly, the range is closed. Indeed, if $(A - z)\phi_n \rightarrow \psi$ then applying (2.12) directly shows that $\|\phi_n\|$ is bounded and so can be replaced by a weakly convergent subsequence. Applying (2.12) again to $\phi_n - \phi_m$ shows that the sequence is actually Cauchy, hence converges to $\phi$ so $(A - z)\phi = \psi$ is in the range. Finally, the orthocomplement to this range is the null space of $A^* - \bar{z}$, which is also trivial, so $A - z$ is an isomorphism and (2.12) also shows that the inverse is bounded, in fact

\begin{equation}
\| (A - z)^{-1} \| \leq \frac{1}{\text{Im } z}.
\end{equation}

When $z \in \mathbb{R}$ we can replace $A$ by $A'$ satisfying (2.8). Then we have to show that $A' - z$ is invertible for $|z| > \|A\|$, but that is shown in the proof of Proposition 2.1.

The basic estimate leading to the spectral theorem is:

**Proposition 2.3.** If $A$ is a bounded self-adjoint operator and $p$ is a real polynomial in one variable,

\begin{equation}
p(t) = \sum_{i=0}^{N} c_i t^i, \ c_N \neq 0,
\end{equation}

then $p(A) = \sum_{i=0}^{N} c_i A^i$ satisfies

\begin{equation}
\| p(A) \| \leq \sup_{t \in [m, M]} |p(t)|.
\end{equation}

**Proof.** Clearly, $p(A)$ is a bounded self-adjoint operator. If $s \notin p([m, M])$ then $p(A) - s$ is invertible. Indeed, the roots of $p(t) - s$ must cannot lie in $[m, M]$, since otherwise $s \in p([m, M])$. Thus, factorizing $p(s) - t$ we have

\begin{equation}
p(t) - s = c_N \prod_{i=1}^{N} (t - t_i(s)), \ t_i(s) \notin [m, M] \implies (p(A) - s)^{-1} \text{ exists}
\end{equation}

since $p(A) = c_N \sum_{i} (A - t_i(s))$ and each of the factors is invertible.

Thus $\text{spec}(p(A)) \subset p([m, M])$, which is an interval (or a point), and from Proposition 2.3 we conclude that $\| p(A) \| \leq \sup p([m, M])$ which is (2.15).

Now, reinterpreting (2.15) we have a linear map

\begin{equation}
P(\mathbb{R}) \ni p \mapsto p(A) \in \mathcal{B}(H)
\end{equation}
from the real polynomials to the bounded self-adjoint operators which is continuous with respect to the supremum norm on \([m, M]\). Since polynomials are dense in continuous functions on finite intervals, we see that (2.17) extends by continuity to a linear map
\[
\mathcal{C}([m, M]) \ni f \mapsto f(A) \in \mathcal{B}(H), \quad \|f(A)\| \leq \|f\|_{[m, M]}, \quad fg(A) = f(A)g(A)
\]
where the multiplicativity follows by continuity together with the fact that it is true for polynomials.

Now, consider any two elements \(\phi, \psi \in H\). Evaluating \(f(A)\) on \(\phi\) and pairing with \(\psi\) gives a linear map
\[
\mathcal{C}([m, M]) \ni f \mapsto \langle f(A)\phi, \psi \rangle \in \mathbb{C}.
\]
This is a linear functional on \(\mathcal{C}([m, M])\) to which we can apply the Riesz representation theorem and conclude that it is defined by integration against a unique Radon measure \(\mu_{\phi, \psi}\):
\[
\langle f(A)\phi, \psi \rangle = \int_{[m,M]} f \, d\mu_{\phi, \psi}.
\]
The total mass \(|\mu_{\phi, \psi}|\) of this measure is the norm of the functional. Since it is a Borel measure, we can take the integral on \((-\infty, b]\) for any \(b \in \mathbb{R}\) ad, with the uniqueness, this shows that we have a continuous sesquilinear map
\[
P_b(\phi, \psi) : H \times H \ni (\phi, \psi) \mapsto \int_{[m,b]} d\mu_{\phi, \psi} \in \mathbb{R}, \quad |P_b(\phi, \psi)| \leq \|A\|\|\phi\||\psi||.
\]
From the Hilbert space Riesz representation theorem it follows that this sesquilinear form defines, and is determined by, a bounded linear operator
\[
P_b(\phi, \psi) = \langle P_b\phi, \psi \rangle, \quad \|P_b\| \leq \|A\|.
\]
In fact, from the functional calculus (the multiplicativity in (2.18)) we see that
\[
P_b^* = P_b, \quad P_b^2 = P_b, \quad \|P_b\| \leq 1,
\]
so \(P_b\) is a projection.

Thus the spectral theorem gives us an increasing (with \(b\)) family of commuting self-adjoint projections such that \(\mu_{\phi, \psi}((-\infty, b]) = \langle P_b\phi, \psi \rangle\) determines the Radon measure for which (2.20) holds. One can go further and think of \(P_b\) itself as determining a measure
\[
\mu((-\infty, b]) = P_b
\]
which takes values in the projections on $H$ and which allows the functions of $A$ to be written as integrals in the form

\[(2.25) \quad f(A) = \int_{[m,M]} f d\mu\]

of which (2.20) becomes the ‘weak form’. To do so one needs to develop the theory of such measures and the corresponding integrals. This is not so hard but I shall not do it.