A Hilbert Space Summary

I will start very briefly, without proofs for the first few results. Generally speaking, if I state here something without a proof it is because I assume that you can prove it on your own. I could be wrong, let me know. I will use the symbol $\mathbb{K}$ to denote the field; that is $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

A Hilbert space is an inner product space over $\mathbb{K}$ that is complete in the metric defined by the inner product. In these notes, the symbol $H$ will always denote a Hilbert space. Basic properties of the inner product are

\[
\begin{align*}
|\langle f, g \rangle| & \leq \|f\| \|g\| \quad \text{(Cauchy-Schwarz inequality)} \\
\|f + g\|^2 & = \|f\|^2 + \|g\|^2 + 2\Re \langle f, g \rangle \quad \text{(Minkowski, or: the norm is a norm)} \\
\|f - g\|^2 & = \|f\|^2 + \|g\|^2 - 2\Re \langle f, g \rangle \\
\|f + g\|^2 + \|f - g\|^2 & = 2\|f\|^2 + 2\|g\|^2 \quad \text{(Parallelogram law)}
\end{align*}
\]

As a metric space, it makes sense to talk of open and closed, and compact, subsets of $H$. And also of convergence.

**Proposition 1** Let $\{f_n\}, \{g_n\}$ be sequences in the Hilbert space $H$ converging to $f, g$, respectively. Then

\[
\begin{align*}
\lim_{n \to \infty} cf_n &= cf \quad \forall c \in \mathbb{K}.
\end{align*}
\]

\[
\begin{align*}
\lim_{n \to \infty} (f_n + g_n) &= f + g.
\end{align*}
\]

\[
\begin{align*}
\lim_{n \to \infty} (f_n, g_n) &= \langle f, g \rangle.
\end{align*}
\]

**Proposition 2** Let $M$ be a finite dimensional subspace of $H$. Then $M$ is closed.

*Proof.* or rather a sketch of a proof. The quickest way is to use a rather nice result that states that in a finite dimensional vector space (over $\mathbb{R}$ or $\mathbb{C}$) all norms are equivalent. That means that our $M$, with the Hilbert space norm, is essentially the same as $\mathbb{R}^d$ or $\mathbb{C}^d$, with $d = \text{dim } M$. In particular it is complete. So if we have a sequence in $M$ that converges to some point in $H$, because it converges it is a Cauchy sequence. Being a Cauchy sequence in the complete space $M$, it must have a limit in $M$. But in a metric space, a sequence can have only one limit, so the original limit has to be in $M$. \(\blacksquare\)

The result is not true if the dimension of the subspace is infinite. For example the set

$$M = \{(a_n)_{n=-\infty}^{\infty} : a_n \neq 0 \text{ for only a finite number of } n\}$$

is a subspace; it is not closed. Clearly $M \neq \ell^2(\mathbb{Z})$, but the closure $\overline{M} = \ell^2(\mathbb{Z})$.

**Proposition 3** Let $M$ be a subspace of the Hilbert space $H$. Then $\overline{M}$, the closure of $M$ in $H$, is also a subspace.

Let $H$ be a Hilbert space. We say $f, g$ are mutually orthogonal, or that $f$ is orthogonal to $g$, and write $f \perp g$, iff $\langle f, g \rangle = 0$. If $A \subseteq H$, we define

$$A^\perp = \{f \in H : f \perp g \text{ for all } g \in A\}.$$

We sometimes (or often) will write $f \perp A$ for $f \in A^\perp$.

**Proposition 4** Assume $A \subseteq H$. Then $A^\perp$ is a closed subspace of $H$. 

We also have the fairly obvious result that doesn’t need to be graced by calling it a proposition:

\[ \|f + g\|^2 = \|f\|^2 + \|g\|^2 \quad \text{if and only if} \quad f \perp g. \]

As a consequence, if \( f_1, \ldots, f_N \) are mutually orthogonal vectors; i.e., \( (f_j, f_k) = 0 \) if \( j \neq k \), then

\[ \left\| \sum_{j=1}^{N} f_j \right\|^2 = \sum_{j=1}^{N} \|f_j\|^2. \]

I’ll interrupt with a nice little exercise; a picture helps. The result will be used soon.

**Exercise 1** Let \( f, g, h \) in \( H \) and assume \( g \neq h \) but \( \|f - g\| = \|f - h\| \). Then

1. \( f - \frac{1}{2}(g + h) \perp g - h \).
2. \( \|f - \frac{1}{2}(g + h)\| < \|f - g\| \).

(In an isosceles triangle, the median to the base is perpendicular to the base and shorter than the two equal sides.)

A subset \( E \) of the Hilbert space \( H \) is said to be an **orthonormal set** iff all elements of \( E \) have norm 1, and any two are mutually orthogonal:

\[ (f, g) = \begin{cases} 0, & \text{if } f \neq g, \\ 1, & \text{if } f = g. \end{cases} \]

**Theorem 5** Let \( C \) be a closed convex non-empty subset of \( H \). If \( f \in H \) there exists a unique \( g \in C \) such that \( \|f - g\| = \text{dist}(f, C) \).

**Proof.** Let \( f \in H \) and let \( d = \text{dist}(f, C) \); since \( C \neq \emptyset \) it follows that \( d \geq 0 \). There is a sequence \( \{g_n\} \) in \( C \) such that

\[ \lim_{n \to \infty} \|f - g_n\| = d. \]

The trick now is to use the following identity:

\[ \|g_n - g_m\|^2 = 2\|f - g_n\|^2 + 4\|f - g_m\|^2 - 4\|f - \frac{1}{2}(g_n + g_m)\|^2. \]

To verify it more or less fast notice that by the parallelogram law

\[
4\|f - \frac{1}{2}(g_n + g_m)\|^2 = 4\|\frac{1}{2}(f - g_n) + \frac{1}{2}(f - g_m)\|^2 = \|(f - g_n) + (f - g_m)\|^2 \\
= \|f - g_n\|^2 + \|f - g_m\|^2 - \|(f - g_n) - (f - g_m)\|^2 \\
= 2\|f - g_n\|^2 + 2\|f - g_m\|^2 - \|g_n - g_m\|^2.
\]

Because \( C \) is convex, \( \frac{1}{2}(g_n + g_m) \in C \) and it follows that \( \|f - \frac{1}{2}(g_n + g_m)\|^2 \geq d^2 \). The tricky identity thus implies

\[ \|g_n - g_m\|^2 \leq 2\|f - g_n\|^2 + 2\|f - g_m\|^2 - 4d^2, \]

from which we get \( \limsup_{m,n \to \infty} \|g_n - g_m\|^2 \leq 2d^2 + 2d^2 - 4d^2 = 0 \); since \( \|g_n - g_m\|^2 \geq 0 \) for all \( m, n \), we conclude \( \lim_{m,n \to \infty} \|g_n - g_m\|^2 = 0 \); that is, the sequence \( \{g_n\} \) is a Cauchy sequence, hence converges. If \( g = \lim_{n \to \infty} g_n \) we gave \( g \in C \) since \( C \) is closed, and \( \|f - g\| = \text{dist}(f, C) \).

Uniqueness is an immediate consequence of Exercise 1.

The main application for us is to the case in which the closed convex subset is a closed subspace \( M \).

**Definition 1** Assume \( M \) is a closed subspace of \( H \). We define a map \( P_M : H \to H \) defining \( g = P_M f \) as the unique element \( g \in M \) such that \( \|f - g\| = \text{dist}(f, M) \).
**Theorem 6** Assume $M$ is a closed subspace of $H$. The following properties hold.

1. $P_M(H) = M$ and $P_M f = f$ for all $f \in M$.
2. $g = P_M f$ if and only if $g \in M$ and $f - g \in M^\perp$.
3. $P_M : H \to M$ is linear.

**Proof.** Property 1. is obvious from the definition. For 2., assume first that $g = P_M f$. By construction $g \in M$, so we must prove $(f - g, h) = 0$ for all $h \in M$. Letting $h \in M$, consider the function

$$
\phi(t) = \|f - (g + th)\|^2
$$

for $t \in \mathbb{R}$. Because $g + th \in M$, this function has a minimum value for $t = 0$, thus

$$
0 = \phi'(0) = \frac{d}{dt}\|f - g - th\|^2\bigg|_{t=0} = \frac{d}{dt}\left(\|f - g\|^2 - 2t\text{Re}(f - g, h) + t^2\|h\|^2\right)\bigg|_{t=0} = -2\text{Re}(f - g, h).
$$

Thus $\text{Re} (f - g, h) = 0$; all that is needed if the space is real. If $\mathbb{K} = \mathbb{C}$, we repeat the argument with $h$ replaced by $ih$ to get $0 = \text{Re} (f - g, ih) = \text{Im}(f - g, h)$.

Conversely, assume $f \in H, g \in M$ and $f - g \perp M$. We then have for all $h \in M$ (since $g - h \in M$ so $f - g \perp g - h$)

$$
\|f - h\|^2 = \|f - g + g - h\|^2 = \|f - g\|^2 + \|g - h\|^2 > \|f - Pf\|^2
$$

except if $h = g$. Because by Theorem 5 the minimum is realized, it must be $g$, so $g = P_M f$. This proves 2.

To see that $P_M$ is linear, let $f, g \in H, a, b \in \mathbb{K}$. By the direct part of 2, $f - PMf, g - PMg \in M^\perp$. But then

$$
af + bg - (aP_M f + bP_M g) = a(f - P_M f) + b(g - P_M g) \in M^\perp,
$$

implying that $aP_M f + bP_M g = PM(af + bg)$. □

The following result is a very basic important Hilbert space result. Its proof is now quite easy.

**Theorem 7** Let $M$ be a closed subspace of $H$. Then $H = M \oplus M^\perp$. In other words, every element $f$ of $H$ can be written uniquely in the form $f = g + h$, where $g \in M$, $h \in M^\perp$.

**Proof.** Let $P = P_M$. If $f \in H$ we thus have $f = g + h$ where $g = Pf \in M$, and $h = (I - P)f = f - Pf \in M^\perp$. This implies that $H = M + M^\perp$. Since $M \cap M^\perp = \{0\}$ is clear ($f \in M \cap M^\perp \Rightarrow (f, F) = 0 \Rightarrow f = 0$), this proves $H = M \oplus M^\perp$. □

Certain immediate useful consequences of all this are summarized in the following proposition. If $A \subseteq H$ we define $A^\perp = (A^\perp)^\perp$.

**Proposition 8**

1. If $M$ is a closed subspace of $H$, then $M^\perp \perp = M$
2. If $A \subseteq H$, then $A^\perp \perp$ is the smallest closed subspace of $H$ containing $A$.
3. A subset $A$ of $H$ spans a dense subspace of $H$ if and only if $f \perp A$ implies that $f = 0$.
4. $P_{M^\perp} = I - P_M$

**Proof.**

1. There could be a shorter proof of this point. It being clear that $M \subseteq M^\perp \perp$, let $f \in M^\perp \perp$; we have to see $f \in M$. By Theorem 7, we can write $f = g + h$, with $g \in M, h \in M^\perp$. Then $(f, h) = 0$ and

$$
\|f\|^2 = (f, g) + (f, h) = (f, g).
$$

By Cauchy-Schwarz,

$$
\|f\|^2 = |(f, g)| \leq \|f\|\|g\|.
$$

This implies that $\|f\| \leq \|g\|$ (Trivial if $\|f\| = 0$; otherwise, cancel). Then, because $(g, h) = 0$,

$$
\|g\|^2 \geq \|f\|^2 = \|g\|^2 + \|h\|^2
$$

implying $\|h\| = 0$. Thus $f = g \in M$. □
2. Let $M$ be the smallest closed subspace containing $A$. In other words, $M$ is the closure of $M_0$, where $M_0$ is the set (space) of all finite combinations of elements of $A$. It is clear that $f \perp A$ implies $f \perp M_0$; hence also $f \perp M$. That is $A^\perp = M^\perp$ thus $A^{\perp \perp} = M^{\perp \perp} = M$.

3. Let $D$ be the subspace spanned by $A$. As mentioned in the proof of the previous point, $A^\perp$ is the same as $D^\perp$ so that $A$ spans a dense subspace if and only if $D = H$, which is if and only if $D^\perp = H^\perp = \{0\}$. We are using here implicitly that the map $M \mapsto M^\perp$ is bijective for closed subspaces; clear since by point 1. it is its own inverse.

4. Let $f \in H$, $g = (I - P_M)f \in M^\perp$ and $P_M f = f - g \in M = (M^\perp)^\perp$. Thus $g = P_M f$.

There are several equivalent ways of defining an orthonormal basis. Our text uses the one that could be easiest to verify, so I’ll stick with it: An orthonormal subset $\mathcal{E}$ of $H$ is an orthonormal basis of $H$ iff $f \in H$, $f \perp \mathcal{E}$ implies $f = 0$. Equivalently if $\mathcal{E}^\perp = \{0\}$. Also equivalently, if $\mathcal{E}$ is a maximal orthonormal set; there is no orthonormal set properly containing $\mathcal{E}$. An orthonormal subset of $H$ is frequently written as an indexed set, say $\mathcal{E} = \{e_\alpha\}_{\alpha \in A}$ and called an orthonormal system.

I will simplify things from now on by following our textbook’s lead and assume that $H$ is separable. The simplification is actually not major. One can see that in this case every orthonormal system has to be countable.

**Proposition 9** Let $\{e_j\}_{j=1}^N$ be a finite orthonormal set. For each $f \in H$:

1. (Bessel) $\sum_{k=1}^N |(f, e_k)|^2 \leq \|f\|^2$.

2. Of all possible choices of coefficients $a_1, \ldots, a_N \in \mathbb{K}$ the one that minimizes $\|f - \sum_{k=1}^N a_k e_k\|$ is $a_k = \langle f, e_k \rangle$, $k = 1, \ldots, N$. In other words $\left\| f - \sum_{k=1}^N \langle f, e_k \rangle e_k \right\| \leq \left\| f - \sum_{k=1}^N a_k e_k \right\|$ for all $a_1, \ldots, a_N \in \mathbb{K}$.

**Proof.** For Bessel’s inequality, we have that

$$0 \leq \left\| f - \sum_{k=1}^N \langle f, e_k \rangle e_k \right\|^2 = \|f\|^2 - 2\Re \left( f, \sum_{k=1}^N \langle f, e_k \rangle e_k \right) + \left\| \sum_{k=1}^N \langle f, e_k \rangle e_k \right\|^2$$

$$= \|f\|^2 - 2\Re \sum_{k=1}^N \langle f, e_k \rangle \langle f, e_k \rangle + \sum_{k=1}^N \|\langle f, e_k \rangle\|^2$$

$$= \|f\|^2 - 2 \sum_{k=1}^N |\langle f, e_k \rangle|^2 + \sum_{k=1}^N |\langle f, e_k \rangle|^2 = \|f\|^2 - \sum_{k=1}^N |\langle f, e_k \rangle|^2$$

Bessel follows. For the minimization property, I will go a bit faster. Let $a_1, \ldots, a_N \in \mathbb{K}$. Notice first that

$$\sum_{k=1}^N |a_k - \langle f, e_k \rangle|^2 = \sum_{k=1}^N |a_k|^2 + \sum_{k=1}^N |\langle f, e_k \rangle|^2 - 2\Re \sum_{k=1}^N a_k \langle f, e_k \rangle.$$ 

or

$$2\Re \sum_{k=1}^N a_k \langle f, e_k \rangle = -\sum_{k=1}^N |a_k - \langle f, e_k \rangle|^2 + \sum_{k=1}^N |a_k|^2 + \sum_{k=1}^N |\langle f, e_k \rangle|^2.$$
Thus
\[
\left\| f - \sum_{k=1}^{N} a_k e_k \right\|^2 = \|f\|^2 - 2\text{Re} \sum_{k=1}^{N} a_k (f, e_k) + \sum_{k=1}^{N} |a_k|^2 \\
= \|f\|^2 - \left(- \sum_{k=1}^{N} |a_k - (f, e_k)|^2 + \sum_{k=1}^{N} |a_k|^2 + \sum_{k=1}^{N} |(f, e_k)|^2 \right) + \sum_{k=1}^{N} |a_k|^2 \\
= \|f\|^2 + \sum_{k=1}^{N} |a_k - (f, e_k)|^2 + \sum_{k=1}^{N} |(f, e_k)|^2.
\]

Clearly, the last expression is smallest possible when \(a_k = (f, e_k)\) for \(k = 1, \ldots, N\).

From now on we will mostly assume that our orthonormal systems are infinite countable. As a corollary of Bessel’s inequality we get

**Corollary 10** Let \(E = \{e_k\}_{k=1}^{\infty}\) be an orthonormal system. For every \(f \in H\), we have

\[
\sum_{k=1}^{\infty} |(f, e_k)|^2 \leq \|f\|^2 < \infty.
\]

We now can prove all the equivalences of being an orthonormal basis.

**Theorem 11** Let \(E = \{e_k\}_{k \in \mathbb{N}}\) be an orthonormal system. The following statements are equivalent.

1. \(E\) is an orthonormal basis.

2. \(E\) spans a dense subset; that is, if we define \(M\) to be the set of all finite linear combinations of elements of \(E\); namely,

\[
M = \{ \sum_{k=1}^{N} c_k e_k : N \in \mathbb{N}, c_1, \ldots, c_N \in \mathbb{K} \},
\]

then \(M\) is dense in \(H\).

3. For each \(f \in H\),

\[
f = \sum_{k=1}^{\infty} (f, e_k) e_k.
\]

That is, \(\lim_{N \to \infty} \left\| f - \sum_{k=1}^{N} (f, e_k) e_k \right\| = 0\).

4. Parseval holds for all elements of \(H\); that is

\[
\|f\|^2 = \sum_{k=1}^{\infty} |(f, e_k)|^2
\]

for all \(f \in H\).

**Proof.** 1. \(\Rightarrow\) 2. Assume 1. so \(E^\perp = \{0\}\). That \(E\) spans a dense subspace is immediate from Proposition 8.

2. \(\Rightarrow\) 3. Assume 2. Let \(f \in H\). Let \(\epsilon > 0\). Since \(\sum_{k} |(f, e_k)|^2 < \infty\) by Corollary 10, there is \(N\) such that

\[
\sum_{k=N+1}^{\infty} |(f, e_k)|^2 < \epsilon^2 / 4.
\]

By assumption 2., there is a finite linear combination of elements of \(E\) at distance \(< \epsilon/2\) from \(f\); by adding zero coefficients if needed we may assume it is of the form \(\sum_{k=1}^{M} a_k e_k\), with \(a_1, \ldots, a_M \in \mathbb{K}\); that is,

\[
\|f - \sum_{k=1}^{M} a_k e_k\| < \epsilon / 2.
\]
Adding even more 0 coefficients, we can assume $M \geq N$. By Proposition 9, part 2, we then have

$$\|f - \sum_{k=1}^{M} (f, e_k)e_k\| < \epsilon/2.$$ 

It follows that if $n \geq \max(N, M)$,

$$\|f - \sum_{k=1}^{n} (f, e_k)e_k\| \leq \|f - \sum_{k=1}^{M} (f, e_k)e_k\| + \|\sum_{k=M+1}^{n} (f, e_k)e_k\| < \epsilon/2 + \left(\sum_{k=M+1}^{n} \|f, e_k\|^2\right)^{1/2} = \epsilon/2 + \epsilon = \epsilon.$$

3. $\Rightarrow$ 4. If $f \in H$, then by 3., $f = \lim_{N \to \infty} \sum_{k=1}^{N} (f, e_k)e_k$ so that

$$\|f\|^2 = (f, f) = \lim_{N \to \infty} \sum_{k=1}^{N} (f, e_k)e_k = \lim_{N \to \infty} \sum_{k=1}^{N} \|f, e_k\|^2 = \sum_{k=1}^{\infty} \|f, e_k\|^2.$$

3. $\Rightarrow$ 1. is immediate. 

Because we assume $H$ is separable, it is not to hard to see it has a countable orthonormal basis. In fact, let $D$ be a dense subset. Let’s order the elements of $D$ as a sequence: $D = \{f_1, f_2, \ldots\}$. Construct a subsequence $\{f_{n_k}\}$ as follows. $f_{n_1}$ is the first non-zero element of the sequence; if there is none then $H = \{0\}$ and we are done. Assume now $f_{n_1}, \ldots, f_{n_k}$ selected for some $k \geq 1$; $1 \leq n_1 < \cdots < n_k$. If these vectors span $H$, we are done. Stop. If not, let $n_{k+1}$ be the first integer $> n_k$ such that $f_{n_{k+1}}$ is linearly independent from $f_{n_1}, \ldots, f_{n_k}$. This process produces a finite or infinite sequence $\{f_{n_k}\}_{k \in A}$, where $A = \mathbb{N}$ or $A = \{1, \ldots, m\}$. The construction guarantees that it is a linearly independent set; all finite combinations of its elements are non-zero, except if all coefficients are zero. It also guarantees that it spans a dense subspace because everything originally in $D$ is in the span. In fact, if $f \in D$ then $f = f_n$ for some $n \in \mathbb{N}$. If $n = n_k$ for some $k$, then it is obviously in the span of the subsequence (being in the subsequence). If not, there is a last $n_k < n$. Since $n$ is not $n_{k+1}$, this means that $f_n$ is a linear combination of $f_{n_1}, \ldots, f_{n_k}$.

The next step is to apply the Gram-Schmidt procedure. In other words, we define $e_1, e_2, \ldots$ by $e_1 = \frac{1}{\|f_{n_1}\|}f_{n_1}$ and, assuming $e_1, \ldots, e_k$ selected (and $k < m$ in the finite case), we define first

$$g_{k+1} = f_{n_{k+1}} - \sum_{j=1}^{k} (f_{n_{k+1}}, e_j)e_j;$$

verify that $g_{k+1} \perp \{e_1, \ldots, e_k\}$; $g_{k+1} \neq 0$ and finally set $e_{k+1} = \frac{1}{\|g_{k+1}\|}g_{k+1}$. We now have an orthonormal system $\{e_k\}_{k \in A}$ and because the span of $\{e_1, \ldots, e_j\}$ is the same as the span of $\{f_1, \ldots, f_j\}$ for all $j$, the system spans the same subspace as $\{f_{n_k}\}_{k \in A}$; that is, it spans a dense subspace. By Theorem 11 we have an orthonormal basis, also known as mons$^*$. 

If $V, W$ are finite dimensional Hilbert spaces, all linear mappings from $V$ to $W$ are continuous. This is false once infinite dimensions are allowed. Here is a simple result that is actually valid for all normed spaces (no completeness needed); I include a proof.

**Theorem 12** Let $H, K$ be Hilbert spaces (over the same field) and assume that $T : H \to K$ is linear. The following properties are equivalent:

1. $T$ is continuous.
2. $T$ is continuous at 0.

$^*$Maximal orthonormal system.
3. $T$ is bounded, meaning there exists a constant $C \geq 0$ such that $\|Tf\| \leq C\|f\|$ for all $f \in H$. (We are using the same symbol for the norm in $H$ as in $K$. With a bit of luck this won’t cause any confusion.)

Proof. 1 $\Rightarrow$ 2 is obvious.

To prove 2 $\Rightarrow$ 3, assume 2. Taking the proverbial $\epsilon$ in the definition of continuity equal to 1, we see that there is $\eta > 0$ such that $\|f\| \leq \eta$ implies $\|Tf\| \leq 1$. Assume now $f \in H$, $f \neq 0$ Then $g = \frac{\eta}{\|f\|}f$ satisfies $\|g\| = \eta \leq \eta$; thus $\|Tg\| \leq 1$, which works out to $\|Tf\| \leq \frac{1}{\eta} \|f\|$ Taking $C = 1/\eta$ we proved $\|Tf\| \leq C\|f\|$ for all $f \in H$, $f \neq 0$; it being trivially true if $f = 0$, we proved it for all $f \in H$, proving 3.

To prove 3 $\Rightarrow$ 1, assume 3. Let $f \in H$. Let $\epsilon > 0$ be given. Let $\delta = \epsilon/C$, where $C$ is the constant of the statement of 3. If $\|g - f\| < \delta$, then

$$\|Tg - Tf\| = \|T(g - f)\| \leq C\|g - f\| < C\delta = \epsilon.$$ 

Continuity follows.

If $T : H_1 \to H_2$ is linear and bounded, one defines the norm of $T$ by

$$\|T\| = \inf \{C : \|Tf\| \leq C\|f\| \ f \in H_1\}.$$ 

It is easy to see that the set of all bounded operators from $H_1$ to $H_2$, sometimes denoted by $B(H_1, H_2)$ is vector space over $K$ with the obvious operations and $\| \cdot \|$ is a norm, meaning

1. $\|T\|g_0 = 0$ for all $T \in B(H_1, H_2)$, and $\|T\| = 0$ if and only if $T$ is the zero operator.
2. $\|cT\| = |c|\|T\|$ for all $c \in K$, $T \in B(H_1, H_2)$.
3. $\|T + S\| \leq \|T\| + \|S\|$ for all $T, S \in B(H_1, H_2)$.

Assume now $g \in H$. We can define a linear map $\ell_g$ from $H \to K$ by $\ell_g(f) = \langle f, g \rangle$. This map is bounded;

$$\|\ell_g(f)\| = \|\langle f, g \rangle\| \leq \|g\| \|f\|.$$ 

The inequality proves $\|\ell_g\| \leq \|g\|$; because $\|\ell_g(g)\| = \|g\|^2$, it follows that $\|\ell_g\| = \|g\|$. The Riesz representation theorem states that these are the only bounded linear functionals (i.e., maps to the field) in $H$:

**Theorem 13** Riesz Representation Assume $\ell : H \to K$ is a bounded linear functional; that is, $\ell$ is linear and there exists a constant $C$ such that $|\ell(f)| \leq C\|f\|$ for all $f \in H$. There exists then a unique $g \in H$ such that $\ell = \ell_g$; i.e., $\ell(f) = \langle f, g \rangle$ for all $f \in H$.

Proof. If $\ell = 0$ (maps everything to 0 in $K$), then $g = 0$, and we are done. So assume $g \neq 0$ and let

$$M = \ker \ell = \{f \in H : \ell(f) = 0\}.$$ 

Because $\ell \neq 0$, $M \neq H$, hence $\perp \neq \{0\}$. There is thus $g_0 \in M^\perp$ such that $c = \ell(g_0) \neq 0$. Let $f \in H$, then

$$\ell(f - \frac{\ell(f)}{c} g_0) = \ell(f) - \frac{\ell(f)}{c} \ell(g_0) = \ell(f) - \ell(f) = 0,$$

so that $h = f - \frac{\ell(f)}{c} g_0 \in M$ and $f = h + \frac{\ell(f)}{c} g_0$ must be the decomposition of $f$ into an element of $M$ plus one of $M^\perp$. Since $(h, g_0) = 0$ we get taking inner product with $g_0$ that

$$0 = \langle f, g_0 \rangle - \frac{\ell(f)}{c} \langle g_0, g_0 \rangle,$$

which can be solved to $\ell(f) = \langle f, g \rangle$ where $g = \frac{1}{c \|g_0\|^2} g_0$. ■
Appendix

If $H$ is separable, then every orthonormal system is countable.

Proof. Assume $E = \{e_\alpha \}_{\alpha \in A}$ is an orthonormal system in the separable Hilbert space $H$. Let $D$ be a dense subset. For each $f \in D$ let $A_f = \{ \alpha \in A : (f, e_\alpha) \neq 0 \}$. The set of indices $A_f$ is countable. In fact, one can prove essentially as done in the text or the notes that

$$\sum_{\alpha \in A} |(f, e_\alpha)|^2 \leq \|f\|^2 < \infty,$$

and this implies that the set of non-zero terms in that sum is countable. Let $B = \bigcup_{f \in D} A_f$. Then $B$ is countable. We claim $B = A$. In fact, let $\alpha \notin B$. Then $(f, e_\alpha) = 0$ for all $f \in D$. By the density of $D$ there is a sequence $\{f_n\}$ converging to $e_\alpha$. But then

$$1 = (e_\alpha, e_\alpha) = \lim_{n \to \infty} (f_n, e_\alpha) = 0,$$

a clear contradiction. We are done. $\blacksquare$