# Notes on Hilbert Sapce for Math 501

Eric A. Carlen[1]
Rutgers University

November 23, 2014

## 1 Introduction

Let $(X, \mathcal{M}, \mu)$ be a measure space. As a set, $L^2(X, \mathcal{M}, \mu)$ consist of the equivalence classes, under equivalence almost everywhere with respect to $\mu$, of functions on $X$ that are $\mathcal{M}$-measurable and such that

$$\int_X |f|^2 \mathrm{d}\mu < \infty \ .$$

For $f \in L^2(X, \mathcal{M}, \mu)$, we define

$$\|f\|_2 := \left( \int_X |f|^2 \mathrm{d}\mu \right)^{1/2} \ . \tag{1.1}$$

Clearly, if $z \in \mathbb{C}$ and $f \in L^2(X, \mathcal{M}, \mu)$, $|zf|^2 = |z|^2|f|^2$ is integrable, and if $f, g \in L^2(X, \mathcal{M}, \mu)$,

$$|f + g|^2 \leq (|f| + |g|)^2 \leq 2(|f|^2 + |g|^2) \ . \tag{1.2}$$

is integrable. Thus, $L^2(X, \mathcal{M}, \mu)$ is a vector space under the usual rules of addition and scalar multiplication for functions.

We shall soon equip $L^2(X, \mathcal{M}, \mu)$ with a metric topology in which it is complete. The next theorem plays a key role in this.

**1.1 THEOREM** (Cauchy-Schwarz inequality for $L^2(X, \mathcal{M}, \mu)$)**.** *Let $f, g \in L^2(X, \mathcal{M}, \mu)$. Then $fg$ is integrable, and*

$$\left| \int_X fg \mathrm{d}\mu \right| \leq \|f\|_2 \|g\|_2 \ . \tag{1.3}$$

*Moreover, there is equality in (1.3) is and only if*

$$\|f\|_2 g = e^{i\theta} \|g\|_2 f \tag{1.4}$$

*where $\theta \in [0, 2\pi)$ is such that*

$$e^{-i\theta} \int_X \overline{f} g \mathrm{d}\mu \geq 0 \ , \tag{1.5}$$

*with $\overline{f}$ denoting the complex conjugate of $f$.*

*Proof.* If either $\|f\|_2 = 0$ or $\|g\|_2 = 0$, the claim is trivially true, so let us suppose that both $\|f\|_2$ and $\|g\|_2$ are strictly positive. Define functions $u$ and $v$ by

$$u = \frac{e^{i\theta}}{\|f\|_2} f \qquad \text{and} \qquad v = \frac{1}{\|g\|_2} g \ ,$$

where $\theta$ is given in (1.5).

Then $\|u\|_2 = \|v\|_2 = 1$, and

$$\|u - v\|_2^2 = \int_X |u|^2 \mathrm{d}\mu + \int_X |v|^2 \mathrm{d}\mu - \int_X \overline{u} v \mathrm{d}\mu - \int_X \overline{v} v \mathrm{d}\mu \ .$$

However, by the choice of $\theta$,

$$\begin{aligned} \int_X \overline{u} v \mathrm{d}\mu &= \frac{1}{\|f\|_2 \|g\|_2} e^{-i\theta} \int_X \overline{f} g \mathrm{d}\mu \\ &= \frac{1}{\|f\|_2 \|g\|_2} \left| \int_X \overline{f} g \mathrm{d}\mu \right| \\ &= \frac{1}{\|f\|_2 \|g\|_2} e^{i\theta} \int_X \overline{g} f \mathrm{d}\mu = \int_X \overline{v} v \mathrm{d}\mu \ . \end{aligned}$$

Therefore,

$$\|u - v\|_2^2 = 2 - 2 \frac{1}{\|f\|_2 \|g\|_2} \left| \int_X \overline{f} g \mathrm{d}\mu \right| \ ,$$

or equivalently,

$$\left| \int_X \overline{f} g \mathrm{d}\mu \right| = \|f\|_2 \|g\|_2 - \frac{1}{2} \|f\|_2 \|g\|_2 \|u - v\|_2^2 \ .$$

From this identity, it follows that when there is equality in (1.3), $u = v$, and thus that (1.4) is satisfied. □

**1.2 DEFINITION** (The $L^2$ inner product)**.** For $f, g \in L^2(X, \mathcal{M}, \mu)$, we define

$$\langle f, g \rangle = \int_X \overline{f} g \mathrm{d}\mu \ . \tag{1.6}$$

Note that for fixed $f \in L^2(X, \mathcal{M}, \mu)$, $g \mapsto \langle f, g \rangle$ is a linear functional on $L^2(X, \mathcal{M}, \mu)$, while for fixed $g \in L^2(X, \mathcal{M}, \mu)$, $f \mapsto \langle f, g \rangle$ is a conjugate linear functional on $L^2(X, \mathcal{M}, \mu)$. The sesquilinear form $\langle \cdot, \rangle$ on $L^2(X, \mathcal{M}, \mu)$ is called the *inner product* on $L^2(X, \mathcal{M}, \mu)$.

The Cauchy-Scwarz inequality can be economically expressed in terms of the inner product as follows:

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2 \ . \tag{1.7}$$

It is also worth noting that

$$\|f\|_2^2 = \langle f, f \rangle \ . \tag{1.8}$$

We are now ready to prove a sharper version of (1.2).

**1.3 THEOREM** (Minkowski Inequality for $L^2(X, \mathcal{M}, \mu)$)**.** *For all* $f, g \in L^2(X, \mathcal{M}, \mu)$,

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2 \ . \tag{1.9}$$

*Proof.*

$$\begin{aligned}\|f+g\|_2^2 &= \langle f+g, f+g\rangle = \|f\|_2^2 + \|g\|_2^2 + \langle f,g\rangle + \langle g,f\rangle \\ &\le \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2 = (\|f\|_2 + \|g\|_2)^2 \ .\end{aligned}$$

□

Therefore, if we define the function $d_2(f,g)$ on $L^2(X,\mathcal{M},\mu)\times L^2(X,\mathcal{M},\mu)$ by

$$d_2(f,g) = \|f-g\|_2 \ ,$$

it follows that for all $f,g,h \in L^2(X,\mathcal{M},\mu)$,

$$d_2(f,h) = \|f-h\|_2 = \|(f-g)+(g-h)\|_2 \le \|f-g\|_2 + \|g-h\|_2 = d_2(f,g) + d_2(g,h) \ .$$

Thus, $f,g \mapsto d_2(f,g)$ satisfies the *triangle inequality*. Since it is also evident that for all $f,g \in L^2(X,\mathcal{M},\mu)$, $d_2(f,g) = d_2(g,f)$ and that $d_2(f,g) = 0$ if and only if $f = g$, $d_2$ is a metric on $L^2(X,\mathcal{M},\mu)$. It is called the $L^2$ *metric.*

The following theorem is fundamental, but will also be familiar given the corresponding theorem that we have proved for $L^1(X,\mathcal{M},\mu)$.

**1.4 THEOREM** (Riesz-Fischer Theorem)**.** *$L^2(X,\mathcal{M},\mu)$ equipped with the $L^2$ metric is complete. Moreover, if $\{f_n\}_{n\in\mathbb{N}}$ is any Cauchy sequence in $L^2(X,\mathcal{M},\mu)$, then there is a subsequence of $\{f_n\}_{n\in\mathbb{N}}$ that converges almost everywhere with respect to $\mu$.*

*Proof.* Let $\{f_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in $L^2(X,\mathcal{M},\mu)$. Recursively define an increasing sequence of number $\{n_k\}_{k\in\mathbb{N}}$ such that

$$\|f_n - f_{n_k}\|_2 \le 2^{-k} \qquad \text{for all} \quad n \ge n_k \ .$$

Since $\{n_k\}_{k\in\mathbb{N}}$ is increasing, it follows that

$$\|f_{n_{k+1}} - f_{n_k}\|_2 \le 2^{-k} \qquad \text{for all} \quad k \ .$$

Now define

$$F_m = |f_{n_1}| + \sum_{k=1}^{m-1} |f_{n_k} - f_{n_{k-1}}| \ .$$

By Theorem 1.3, applied iteratively,

$$\|F_m\|_2 \le \|f_{n_1}\|_2 + \sum_{k=1}^{m-1} \|f_{n_k} - f_{n_{k-1}}\|_2 \le \|f_{n_1}\|_2 + 1 \ .$$

Thus, by the Lebesgue Monotone Convergence Theorem,

$$F := \lim_{m\to\infty} F_m$$

is square-integrable and

$$\int_X F^2 \mathrm{d}\mu \le \|f_{n_1}\|_2 + 1 \ .$$

It follows that $F < \infty$ a.e. $\mu$, and thus that

$$\sum_{k=1}^{\infty}(f_{n_k} - f_{n_{k-1}})$$

is absolutely convergent a.e. $\mu$. But since absolute convergence implies convergence,

$$\lim_{m\to\infty}\left[f_{n_1} + \sum_{k=1}^{m-1}(f_{n_k} - f_{n_{k-1}})\right] = \lim_{m\to\infty} f_{n_m}$$

exists almost everywhere. Call this limit $f$. As a pointwise limit of measurable functions $f$ is measurable. Also, $f \in L^2(X, \mathcal{M}, \mu)$ by Fatou's Lemma.

Next, $|f_{n_m} - f|^2 \leq 4F^2$, and since $4F^2$ is integrable, the Lebesgue Dominated Convergence Theorem implies that

$$\lim_{m\to\infty} \|f_{n_m} - f\|_2 = 0 \ .$$

Thus, a subsequence of the Cauchy sequence $\{f_n\}_{n\in\mathbb{N}}$ converges of $f$ in the $L^2$ metric. But then the whole sequence converges to the same limit. We have found above a subsequence $\{f_{n_m}\}_{m\in\mathbb{N}}$ that converges to $f$ a.e. $\mu$. □

# 2 Hilbert Space

Let $\mathcal{H}$ be a complex vector space. A *sesquilinear form* on $\mathcal{H}$ is a functions on $\mathcal{H}\times\mathcal{H}$ with values in $\mathbb{C}$ such that for fixed $f \in \mathcal{H}$, $g \mapsto \langle f, g\rangle$ is a linear functional on $\mathcal{H}$, while for fixed $g \in \mathcal{H}$, $f \mapsto \langle f, g\rangle$ is a conjugate linear functional on $\mathcal{H}$. The sesquilinear form is *positive definite* in case $\langle f, f\rangle > 0$ for all $f \neq 0$.

**2.1 DEFINITION** (Inner product space)**.** An *inner product space* is a complex vector space $\mathcal{H}$ equipped with a positive definite sesquilinear form $\langle\cdot,\cdot\rangle$ on $\mathcal{H}\times\mathcal{H}$. We define the *norm* $\|f\|$ of a vector $f \in \mathcal{H}$ by

$$\|f\| = \langle f, f\rangle \ . \tag{2.1}$$

The example behind these definitions is $L^2(X, \mathcal{M}, \mu)$ equipped with the sesquilinear form (1.6). The Cauchy-Schwarz inequality holds in this more abstract setting:

If $f, g \in \mathcal{H}$ and $\|f\| \neq 0$ and $\|g\| \neq 0$, let $\theta \in [0, 2\pi)$ be such that

$$e^{-i\theta}\langle f, g\rangle > 0 \ .$$

Define $u = e^{i\theta}\|f\|^{-1}f$ and $v = \|g\|^{-1}g$. Then computations identical to those in the proof of Theorem 1.1 show that

$$|\langle f, g\rangle| = \|f\|_2\|g\|_2 - \frac{1}{2}\|f\|_2\|g\|_2\|u - v\|_2^2 \ .$$

In particular

$$|\langle f, g\rangle| \leq \|f\|_2\|g\|_2 \ . \tag{2.2}$$

Next, computations identical to those in the proof of Theorem 1.2 show that for all $f, g \in \mathcal{H}$,

$$\|f + g\| \leq \|f\| + \|g\| \ . \tag{2.3}$$

It follows that $d(f,g) = \|f-g\|$ is a metric on $\mathcal{H}$. This metric is called the *inner-product metric* on $\mathcal{H}$.

There are two important identities that hold in any complex inner product space $\mathcal{H}$: The *parallelogram identity* is

$$\left\|\frac{f+g}{2}\right\|^2 + \left\|\frac{f-g}{2}\right\|^2 = \frac{\|f\|^2+\|g\|^2}{2} \ . \tag{2.4}$$

The *polarization identity* is

$$\begin{aligned}\langle f,g\rangle &= \frac{1}{4}[\langle f+g,f+g\rangle - \langle f-g,f-g\rangle - i\langle f+ig,f+ig\rangle + i\langle f-ig,f-ig\rangle] \\ &= \frac{1}{4}[\|f+g\|^2 - \|f-g\|^2 - i\|f+ig\|^3 + i\|f-ig\|^2] \ . \end{aligned} \tag{2.5}$$

The polarization identity shows that the correspondence between inner products and norms is one-to-one: Every inner product defines a norm, and the inner product may be recovered from the norm.

**2.2 DEFINITION** (Hilbert Space)**.** A *Hilbert space* is a complex vector space $\mathcal{H}$ equipped with a sesquilinear form $\langle\cdot,\cdot\rangle$ such that $\mathcal{H}$ is complete in its inner product metric.

We have seen that associated to any measure space $(X,\mathcal{M},\mu)$ there is a natural inner product on $\mathcal{H} := L^2(X,\mathcal{M},\mu)$ making $\mathcal{H}$ a Hilbert space. If one takes $X = \{1,\dots,n\}$, $\mathcal{M} = 2^X$, and $\mu$ to be counting measure, $L^2(X,\mathcal{M},\mu) = \mathbb{C}^n$ with its usual inner product structure. Before the invention of the Lebesgue integral, such finite dimensional Hilbert spaces were all that were known. Infinite dimensional inner product spaces based on the Riemann integral were known, but these were not complete. Our next theorem is often invoked in arguments that turn on the completeness of Hilbert space.

**2.3 THEOREM** (Projection Lemma)**.** *Let $K$ be a closed convex set in a Hilbert space. Then $K$ contains a unique element of minimal norm. That is, there exists $v \in K$ such that $\|v\| < \|w\|$ for all $w \in K$, $w \neq v$.*

*Proof.* Let $D := \inf\{\|w\| \ : \ w \in K \ \}$. If $D = 0$, then $0 \in K$ since $K$ is closed, and this is the unique element of minimal norm. Hence we may suppose that $D > 0$. Let $\{w_n\}_{n\in\mathbb{N}}$ be a sequence in $K$ such that $\lim_{n\to\infty}\|w_n\| = D$. By the parallelogram identity

$$\left\|\frac{w_m+w_n}{2}\right\|^2 + \left\|\frac{w_m-w_n}{2}\right\|^2 = \frac{\|w_m\|^2+\|w_n\|^2}{2} \ .$$

By the convexity of $K$, and the definition of $D$, $\left\|\frac{w_m+w_n}{2}\right\|^2 \geq D^2$ and so

$$\left\|\frac{w_m-w_n}{2}\right\|^2 = \frac{\left(\|w_m\|^2 - D^2\right) + \left(\|w_n\|^2 - D^2\right)}{2} \ .$$

By construction, the right side tends to zero, and so $\{w_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. Then, by the completeness that is a defining property of Hilbert spaces, $\{w_n\}_{n\in\mathbb{N}}$ is a convergent sequence. Let $v$ denote the limit. By the continuity of the norm, $\|v\| = \lim_{n\to\infty}\|w_n\| = D$. Finally, if $u$ is any other vector in $K$ with $\|u\| = D$, $(u+v)/2 \in K$, so that $\|(u+v)/2\| \geq D$. Then by the parallelogram identity once more $\|(u+v)/2\| = 0$, and so $u = v$. This proves the uniqueness. □

Next, let $\mathcal{H}$ be a Hilbert space and let $L$ be a continuous linear functional on $\mathcal{H}$. Then

$$L^{-1}(\{z \in \mathbb{C} \ : \ |z| < 1\}$$

is open and contains 0, and hence for some $r > 0$, $L^{-1}(\{z \in \mathbb{C} \ : \ |z| < 1\}$ contains $\{f \in \mathcal{H} \ : \ \|f\| < r\}$. Thus, For all $u \in \mathcal{H}$ with $\|u\| < 1$, $|L(u)| = r^{-1}|L(ru)| < 1$. Therefore, the quantity $\|L\|$ defined by

$$\|L\| = \sup\{|L(u)| \ : \ \|u\| \leq 1\} \tag{2.6}$$

satisfies $\|L\| \leq r^{-1} < \infty$.

The quantity $\|L\|$ is called the *norm* of $L$. Any linear functional on $\mathcal{H}$ with finite norm is called a *bounded linear functional*. Note that it is the restriction to the unit ball about the origin that is bounded, not the linear functional on all of $\mathcal{H}$.

For any $f \neq 0$ in $\mathcal{H}$, we may define $u = \|f\|^{-1}f$, so that $f = \|f\|u$ where $\|u\| = 1$. Then since $L$ is linear,

$$|L(f)| = |L(\|f\|u)| = |L(u)|\|f\| \leq \|L\|\|f\|.$$

Conversely, suppose that $L$ is an bounded linear functional on $\mathcal{H}$. Then for any $f, g \in \mathcal{H}$, since $L$ is linear,

$$|L(f) - L(g)| = |L(f - g)| \leq \|L\|\|f - g\| \ ,$$

and so $L$ is Lipschitz continuous on $\mathcal{H}$ with Lipschitz constant $\|L\|$.

The set of continuous linear functional on $\mathcal{H}$ is a vector space known as the dual space to $\mathcal{H}$, which is usually denoted $\mathcal{H}^*$. If $L_1$ and $L_2$ are two elements of $\mathcal{H}^*$, then for any $f \in \mathcal{H}$

$$|(L_1 + L_2)(f)| = |L_1(f) + L_2(f)| \leq |L_1(f)| + |L_2(f)| \ .$$

Therefore

$$\|L_1 + L_2\| \leq \|L_1\| + \|L_2\| \ ,$$

and it readily follows that $d(L_1, L_2) := \|L_1 - L_2\|$ is a metric on $\mathcal{H}^*$. We equip $\mathcal{H}^*$ with this metric.

As an example of a bounded linear functional on $\mathcal{H}$, consider any $v \in \mathcal{H}$, and define

$$L_v(f) = \langle v, f\rangle \ .$$

Then $L_v$ is linear, and by the Cauchy-Schwarz inequality, $|L_v(f)| \leq \|v\|\|f\|$, so that $\|L\| \leq \|v\|$. Thus, $L_v$ is a bounded linear functional on $\mathcal{H}$. The next theorem says that this is the only sort of example.

**2.4 THEOREM** (Riesz Representation Theorem)**.** *Let $L$ be bounded linear transformation on $\mathcal{H}$. Then there is a unique vector$v_L \in \mathcal{H}$ such that*

$$L(f) = \langle v_L, f\rangle$$

*for all $f \in \mathcal{H}$, and moreover, $\|v_L\| = \|L\|$.*

*Proof.* If $L(f) = 0$ for all $f \in \mathcal{H}$, the assertion is trivial, so suppose that $\|L\| > 0$. Define $K$ to be the set

$$K := \{f \in \mathcal{H} \ : \ \Re(L(f)) = \|L\| \ \} \ .$$

It is readily checked that this is aclosed convex set in $\mathcal{H}$.

If $f \in K$, the $\|L\|\|f\| \geq |L(f)| \geq \Re(L(f)) = \|L\|$, and hence $\|f\| \geq 1$. On the other hand, by the definition of $\|L\|$, there is a sequence of unit vectors $\{u_n\}_{n\in\mathbb{N}}$ such that $|L(u_n)| \to \|L\|$, and we may assume $|L(u_n)| > 0$ for all $n$. Then choosing $\theta_n \in [0, 2\pi)$ so that $e^{i\theta_n} L(u_n) = |L(u_n)|$, we have that

$$v_n := e^{i\theta_n} \frac{\|L\|}{\|L(u_n)\|} u_n \in K$$

and $\|v_n\| \to 1$. Thus,

$$\inf\{\|v\| \ : \ v \in K\} = 1 \ .$$

It now follows from the Projection Lemma that there is a unique unit vector $u_0 \in K$. That is, we have found a unit vector $u_0$ such that

$$\|L\| = |L(u_0)| \ . \tag{2.7}$$

Moreover, $L(u_0) = |L(u_)|$ so that $L(u_0) = \|L\|$. The crucial use of the Projection Lemma was to ensure the existence of a unit vector $u_0$ satisfying (2.7). Once this is achieved, by replacing $u_0$ with $e^{i\theta} u_0$, one can achieve $L(u_0) = |L(u_)|$.

Since for all $f$ with $f \neq 0$, $\Re(L(f)) \leq |L(f)| \leq \|L\|\|f\|$,

$$\frac{\Re(L(f))}{\|f\|} \leq \|L\| = \frac{\Re(L(u_0))}{\|u_0\|} \ .$$

Therefore, for any $g \in \mathcal{H}$, the function

$$\varphi(t) = \frac{\Re(L(u_0) + tg)}{\|u_0 + tg\|}$$

is well defined on an open interval about $t = 0$, and has a maximum at $t = 0$. Moreover, one readily checks that $\varphi$ is differentiable there and computes

$$\varphi'(0) = \Re(L(g)) - \|L\|\Re(\langle u_0, g\rangle) \ .$$

Since the left hand side is zero for all $g$, $\Re(L(g)) = \|L\|\Re(\langle u_0, g\rangle)$ for all $g$. Replacing $g$ by $ig$, the same is true of the imaginary parts, and so $L(g) = \langle \|L\| u_0, g\rangle$ for all $g$. Thus, $v_L = \|L\| u_0$ is such that $L(f) = \langle v_L, f\rangle$ for all $f \in \mathcal{H}$, and $\|v_L\| = \|L\|$.

If $w_L$ were any other vector with $L(f) = \langle w_L, f\rangle$ for all $f \in \mathcal{H}$, we would have $\langle v_L - w_L, f\rangle = 0$ for all $f \in \mathcal{H}$, Taking $f = v_L - w_L$, we see that $\|v_L - w_L\|^2 = 0$, and so $w_L = v_L$, proving the uniqueness of $v_L$. □

The Riesz Representation Theorem allows us to identify a Hilbert space $\mathcal{H}$ with its *dual space*; i.e., the space consisting of continuous linear functionals on $\mathcal{H}$. The mapping $L \to v_L$ is an isometry from $\mathcal{H}^*$ onto $\mathcal{H}$. (The range is all of $\mathcal{H}$ by our example just before the theorem.) This mapping is even conjugate linear. Because of this identification, we have used, as is usual, the same notation for the norms on $\mathcal{H}$ and $\mathcal{H}^*$.

# 3 Separability and orthonormal bases

A Hilbert space $\mathcal{H}$ is separable if is contains a countable dense set. We have seen a number of examples of this when $\mathcal{H}$ is of the form$L^2(X, \mathcal{M}, \mu)$ for some measure space $(X, \mathcal{M}, \mu)$ where $\mathcal{M}$ is generated by a countable algebra.

**3.1 DEFINITION** (Complete orthonormal sets )**.** Let $\mathcal{H}$ be a Hilbert space. To vectors $f$ and $g$ in $\mathcal{H}$ are orthogonal in case $\langle f, g\rangle = 0$. A subset of $\mathcal{H}$ is *orthonormal* in case the vectors in it are all unit vectors and are all mutually orthogonal. An orthonormal set $\mathcal{U} \subset \mathcal{H}$ is *complete* if there is no non-zero vector in $\mathcal{H}$ that is orthogonal to every vector in $\mathcal{U}$.

**3.2 THEOREM** (Existence of complete orthonormal sequences)**.** *Let $\mathcal{H}$ be a separable Hilbert space. Then there exists a complete orthonormal sequence $\{u_n\}_{n\in\mathbb{N}}$ in $\mathcal{H}$.*

*Proof.* Let $\{f_n\}_{n\in\mathbb{N}}$ be a dense sequence in $\mathcal{H}$. Apply the Gram-Schmidt orthonormalization procedure to the sequence $\{f_n\}_{n\in\mathbb{N}}$, producing the sequence $\{u_n\}_{n\in\mathbb{N}}$. Now fix any $f \in H$ and suppose that $\langle u_n, f\rangle = 0$ for all $n$, but that $\|f\| > 0$. Then there is an $m$ so that $\|f - f_m\| < \|f\|/2$. Since $f_m$ is a linear combination of $u_1, \dots u_m$ by the nature of the Gram-Schmidt process, $f$ is orthogonal to $f_m$. But then

$$\|f - f_m\|^2 = \langle f - f_m, f - f_m\rangle = \langle f, f\rangle + \langle f_m, f_m\rangle = \|f\|^2 + \|f_m\|^2 \geq \|f\|^2 \ ,$$

and this contradiction proves that there is no such $f$. Thus, the orthonormal set $\{u_n\}_{n\in\mathbb{N}}$ is complete.

□

Next, given $f \in \mathcal{H}$ and an orthonormal sequence $\{u_n\}_{n\in\mathbb{N}}$, let us try to approximate $f$ by finite linear combinations $\sum_{j=1}^n \alpha_j u_j$. There is a unique best choice of the coefficients $\alpha_1, \dots, \alpha_n$: We compute

$$\begin{aligned} \left\| f - \sum_{j=1}^n \alpha_j u_j \right\|^2 &= \left\langle f - \sum_{j=1}^n \alpha_j u_j \ , \ f - \sum_{j=1}^n \alpha_j u_j \right\rangle \\ &= \|f\|^2 - \sum_{j=1}^n 2\Re(\overline{\alpha_j}\langle u_j, f\rangle) + \sum_{j=1} |\alpha_j|^2 \\ &= \|f\|^2 - \sum_{j=1}^n |\langle u_j, f\rangle|^2 + \sum_{j=1}^n |\alpha_j - \langle u_j, f\rangle|^2 \ . \end{aligned} \tag{3.1}$$

It follows that we achieve the best approximation by taking $\alpha_j = \langle u_j, f\rangle$, in which case we have *Bessel's identity*

$$\left\| f - \sum_{j=1}^n \langle u_j, f\rangle u_j \right\|^2 = \|f\|^2 - \sum_{j=1}^n |\langle u_j, f\rangle|^2 \ . \tag{3.2}$$

Since the left hand side is non-negative, this implies *Bessel's inequality*

$$\|f\|^2 \geq \sum_{j=1}^n |\langle u_j, f\rangle|^2 \ . \tag{3.3}$$

Therefore, for all $f \in \mathcal{H}$, the series $\sum_{j=1}^{\infty} |\langle u_j, f\rangle|^2$ converges, and in particular

$$\lim_{m\to\infty} \sum_{j=m+1}^{\infty} |\langle u_j, f\rangle|^2 = 0 \ . \tag{3.4}$$

Now given $f \in \mathcal{H}$ and an orthonormal sequence $\{u_n\}_{n\in\mathbb{N}}$ define

$$g_n := \sum_{j=1}^{n} \langle u_j, f\rangle u_j \ ,$$

so that $\{g_n\}_{n\in\mathbb{N}}$ is the sequence of best approximates to $f$ by linear combinations of more and more terms in our orthonormal sequence.

The sequence $\{g_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. To see this note that for $n \geq m$

$$g_n - g_m = \sum_{j=m+1}^{n} \langle u_j, f\rangle u_j \ ,$$

and so

$$\|g_n = g_m\|^2 = \sum_{j=m+1}^{n} |\langle u_j, f\rangle|^2 \ ,$$

and by (**??**) for all $\epsilon > 0$, $\|g_n = g_m\|^2 \leq \epsilon$ for all $m, n$ sufficiently large.

Since $\mathcal{H}$ is complete, this Cauchy sequence converges to some $g \in \mathcal{H}$. Next notice that for all $n > j$, since $\{u_j\}$ is orthonormal,

$$\langle u_j, g_n\rangle = \langle u_j, f\rangle.$$

Then by continuity of $L_j := \langle u_j, \cdot\rangle$,

$$\langle u_j, g\rangle = \lim_{n\to\infty} \langle u_j, g_n\rangle = \langle u_j, \cdot\rangle \ .$$

Then

$$\langle u_j, g - f\rangle = 0$$

for all $j$. If the orthonormal set is complete, then this entails that $g = f$, and thus that our sequence of successive approximations actually converges to $f$. We summarize:

**3.3 THEOREM.** *Let $\mathcal{H}$ be a Hilbert space that contains a complete orthonormal sequence $\{u_n\}_{n\in\mathbb{N}}$. Then for every $f \in \mathcal{H}$,*

$$f = \sum_{j=1}^{\infty} \langle u_j, f\rangle u_j \ ,$$

*meaning that*

$$\lim_{n\to\infty} \left\| f - \sum_{j=1}^{n} \langle u_j, f\rangle u_j \right\| = 0 \ .$$

*Moreover,*

$$\|f\|^2 = \sum_{j=1}^{\infty} |\langle u_j, f\rangle|^2 \ .$$

*Proof.* It remains to observe that the final statement follows from the first and Bessels' identity. □

Notice that whenever $\mathcal{H}$ contains a complete orthonormal sequence $\{u_j\}_{j\in\mathbb{N}}$, $\mathcal{H}$ is separable since by the theorem rational finite linear combinations of the vectors in $\{u_j\}_{j\in\mathbb{N}}$ are a countable dense set. Hence complete orthonormal sequences will exist in $\mathcal{H}$ if and only if $\mathcal{H}$ is separable.

**3.4 DEFINITION** (Orthonormal basis)**.** Let $\mathcal{H}$ be a separable Hilbert space. A complete orthonormal sequence in $\mathcal{H}$ is called an *orthonormal basis* for $\mathcal{H}$.

**3.5 DEFINITION.** Let $\mathcal{H} = L^2(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ where $\mu$ is counting measure. This Hilbert space is denote $\ell^2$. The elements of $\ell^2$ are the square summable sequences. It is easy to check that if we define $u_j$ to be the sequence whose $j$th term is 1 ,and all other terms are zero, the $\{u_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for $\ell^2$.

Notice that if $\mathcal{H}$ is any separable Hilbert space, and $\{u_j\}_{j\in\mathbb{N}}$ is any orthonormal basis in $\mathcal{H}$, then the transformation that sends $f \in \mathcal{H}$ to the sequence whose $j$th term is $\langle u_j, f\rangle$ is a linear isometry of $\mathcal{H}$ onto $\ell^2$. That is, every separable Hilbert space can be mapped onto $ell^2$ by a linear isometry.

**3.6 EXAMPLE.** *Let $\mathcal{H} = L^2(S^1, \mathcal{B}_{S^1}, m)$ be the Hilbert space of Borel functions $f(\theta)$ on the unit circle that are square integrable with respect to Lebesgue measure $m$ on $S^1$ normalized so that $\mu(S^1) = 1$.*

*It is then readily checked that with $u_j$ defined by*

$$u_n(\theta) = e^{in\theta},$$

*$\{u_n\}_{n\in\mathbb{Z}}$ is orthonormal. By the Stone-Wierstrass Theorem, every continuos function on $S^1$ can be approximated arbitrarily well in the uniform metric by a finite linear combination of the vectors in our orthonormal set. Since continuous functions are dense in $\mathcal{H}$, this is also true of every $f \in \mathcal{H}$. It follows that $\{u_n\}_{n\in\mathbb{Z}}$ is complete: If $f$ is any non-zero vector in $\mathcal{H}$ such that $f$ is orthogonal to every $u_n$, then by (3.1),*

$$\left\| f - \sum_{j=-n}^{n} \alpha_j u_j \right\| \geq \|f\|$$

*no matter how the coefficients $\alpha_{-n} \ldots, \alpha_n$ are chosen. For $\|f\| > 0$ this is impossible by the density of such finite linear combinations. Thus, $\{u_n\}_{n\in\mathbb{Z}}$ is orthonormal basis for $\mathcal{H}$, called the* Fourier basis. *It follows that for each $f$ $in\mathcal{H}$,*

$$f = \lim_{n\to\infty} \sum_{j=-n}^{n} \langle u_j, f\rangle u_j \ .$$

*The sequence $\{\langle u_j, f\rangle\}_{j\in\mathbb{Z}}$ is called the sequence of* Fourier coefficients *of $f$. The associated linear isometry of $\mathcal{H}$ onto $\ell^2$ is the* discrete Fourier transform.

# 4 Excercises

**1.** Let $\mathcal{H}$ be a separable, infinite dimensional Hilbert space, and let $\{u_n\}_{n\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{H}$. Let $\{c_j\}_{\mathrm{J}\in\mathbb{N}}$ be a given sequence of non-negative numbers, and define

Let $C \subset \mathcal{H}$ be defined by

$$C = \{ f \in \mathcal{H} \ : \ \|f\| \leq 1 \quad \text{and} \quad |\langle u_j, f\rangle| \leq c_j \quad \text{for all } j\} \ .$$

Show that $C$ is always closed and bounded, but is compact if and only if $\sum_{j=1}^{\infty} c_j^2 < \infty$. Taking each $c_j = 1$, $C$ becomes the unit ball in $\mathcal{H}$, and thus the unit ball is not compact.

**2.** For real vlaued sqaure integrable functions $f$ on $[-1,1]$, compute

$$\max\{\int_{[-1,1]} x^3 f(x)\mathrm{d}m \ : \ \int_{[-1,1]} x^j f(x)\mathrm{d}m = 0 \quad \text{for} \quad j = 0,1,2 \quad \text{and} \quad \int_{[-1,1]} f^2(x)\mathrm{d}m = 1\}$$

**3.** Show that if $E$ is any Borel set in $(0, 2\pi]$ then

$$\lim_{j\to\infty} \int_E \cos(jx)\mathrm{d}m = \lim_{j\to\infty} \int_E \sin(jx)\mathrm{d}m = 0 \ .$$

Next, consider any incrasing sequence $\{n_k\}$ of the natural numbers. Define $E$ to be the set of all $x$ for which

$$\lim_{k\to\infty} \sin(n_k x) \quad \text{exists} \ .$$

Show that $m(E) = 0$. (The identity $2\sin^2 x = 1 - \cos(2x)$ and the first part may prove useful.)

**4.** Prove the polarization identity (2.5).