4 Hilbert Spaces: An Introduction

> Born barely 10 years ago, the theory of integral equations has attracted wide attention as much as for its inherent interest as for the importance of its applications. Several of its results are already classic, and no one doubts that in a few years every course in analysis will devote a chapter to it.
>
> *M. Plancherel,* 1912

There are two reasons that account for the importance of Hilbert spaces. First, they arise as the natural infinite-dimensional generalizations of Euclidean spaces, and as such, they enjoy the familiar properties of orthogonality, complemented by the important feature of completeness. Second, the theory of Hilbert spaces serves both as a conceptual framework and as a language that formulates some basic arguments in analysis in a more abstract setting.

For us the immediate link with integration theory occurs because of the example of the Lebesgue space $L^2(\mathbb{R}^d)$. The related example of $L^2([-\pi,\pi])$ is what connects Hilbert spaces with Fourier series. The latter Hilbert space can also be used in an elegant way to analyze the boundary behavior of bounded holomorphic functions in the unit disc.

A basic aspect of the theory of Hilbert spaces, as in the familiar finite-dimensional case, is the study of their linear transformations. Given the introductory nature of this chapter, we limit ourselves to rather brief discussions of several classes of such operators: unitary mappings, projections, linear functionals, and compact operators.

1 The Hilbert space L^2

A prime example of a Hilbert space is the collection of **square integrable functions** on $\mathbb{R}^d$, which is denoted by $L^2(\mathbb{R}^d)$, and consists of all complex-valued measurable functions f that satisfy

$$\int_{\mathbb{R}^d} |f(x)|^2\, dx < \infty.$$

The resulting $L^2(\mathbb{R}^d)$**-norm** of f is defined by

$$\|f\|_{L^2(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(x)|^2\, dx\right)^{1/2}.$$

The reader should compare those definitions with these for the space $L^1(\mathbb{R}^d)$ of integrable functions and its norm that were described in Section 2, Chapter 2. A crucial difference is that L^2 has an inner product, which L^1 does not. Some relative inclusion relations between those spaces are taken up in Exercise 5.

The space $L^2(\mathbb{R}^d)$ is naturally equipped with the following inner product:

$$(f, g) = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx, \qquad \text{whenever } f, g \in L^2(\mathbb{R}^d),$$

which is intimately related to the L^2-norm since

$$(f, f)^{1/2} = \|f\|_{L^2(\mathbb{R}^d)}.$$

As in the case of integrable functions, the condition $\|f\|_{L^2(\mathbb{R}^d)} = 0$ only implies $f(x) = 0$ almost everywhere. Therefore, we in fact identify functions that are equal almost everywhere, and define $L^2(\mathbb{R}^d)$ as the space of equivalence classes under this identification. However, in practice it is often convenient to think of elements in $L^2(\mathbb{R}^d)$ as functions, and not as equivalence classes of functions.

For the definition of the inner product (f, g) to be meaningful we need to know that $f\overline{g}$ is integrable on $\mathbb{R}^d$ whenever f and g belong to $L^2(\mathbb{R}^d)$. This and other basic properties of the space of square integrable functions are gathered in the next proposition.

In the rest of this chapter we shall denote the L^2-norm by $\|\cdot\|$ (dropping the subscript $L^2(\mathbb{R}^d)$) unless stated otherwise.

Proposition 1.1 *The space $L^2(\mathbb{R}^d)$ has the following properties:*

(i) *$L^2(\mathbb{R}^d)$ is a vector space.*

(ii) *$f(x)\overline{g(x)}$ is integrable whenever $f, g \in L^2(\mathbb{R}^d)$, and the Cauchy-Schwarz inequality holds: $|(f, g)| \le \|f\|\, \|g\|$.*

(iii) *If $g \in L^2(\mathbb{R}^d)$ is fixed, the map $f \mapsto (f, g)$ is linear in f, and also $(f, g) = \overline{(g, f)}$.*

(iv) *The triangle inequality holds: $\|f + g\| \le \|f\| + \|g\|$.*

Proof. If $f, g \in L^2(\mathbb{R}^d)$, then since $|f(x) + g(x)| \leq 2\max(|f(x)|, |g(x)|)$, we have

$$|f(x) + g(x)|^2 \leq 4(|f(x)|^2 + |g(x)|^2),$$

therefore

$$\int |f + g|^2 \leq 4 \int |f|^2 + 4 \int |g|^2 < \infty,$$

hence $f + g \in L^2(\mathbb{R}^d)$. Also, if $\lambda \in \mathbb{C}$ we clearly have $\lambda f \in L^2(\mathbb{R}^d)$, and part (i) is proved.

To see why $f\overline{g}$ is integrable whenever f and g are in $L^2(\mathbb{R}^d)$, it suffices to recall that for all $A, B \geq 0$, one has $2AB \leq A^2 + B^2$, so that

$$\int |f\overline{g}| \leq \frac{1}{2}\left[\|f\|^2 + \|g\|^2\right]. \tag{1}$$

To prove the Cauchy-Schwarz inequality, we first observe that if either $\|f\| = 0$ or $\|g\| = 0$, then $fg = 0$ is zero almost everywhere, hence $(f, g) = 0$ and the inequality is obvious. Next, if we assume that $\|f\| = \|g\| = 1$, then we get the desired inequality $|(f, g)| \leq 1$. This follows from the fact that $|(f, g)| \leq \int |f\overline{g}|$, and inequality (1). Finally, in the case when both $\|f\|$ and $\|g\|$ are non-zero, we normalize f and g by setting

$$\tilde{f} = f/\|f\| \quad \text{and} \quad \tilde{g} = g/\|g\|,$$

so that $\|\tilde{f}\| = \|\tilde{g}\| = 1$. By our previous observation we then find

$$|(\tilde{f}, \tilde{g})| \leq 1.$$

Multiplying both sides of the above by $\|f\|\,\|g\|$ yields the Cauchy-Schwarz inequality.

Part (iii) follows from the linearity of the integral.

Finally, to prove the triangle inequality, we use the Cauchy-Schwarz inequality as follows:

$$\begin{aligned}
\|f + g\|^2 &= (f + g, f + g) \\
&= \|f\|^2 + (f, g) + (g, f) + \|g\|^2 \\
&\leq \|f\|^2 + 2\,|(f, g)| + \|g\|^2 \\
&\leq \|f\|^2 + 2\,\|f\|\,\|g\| + \|g\|^2 \\
&= (\|f\| + \|g\|)^2,
\end{aligned}$$

and taking square roots completes the argument.

We turn our attention to the notion of a limit in the space $L^2(\mathbb{R}^d)$. The norm on L^2 induces a metric d as follows: if $f, g \in L^2(\mathbb{R}^d)$, then

$$d(f, g) = \|f - g\|_{L^2(\mathbb{R}^d)}.$$

A sequence $\{f_n\} \subset L^2(\mathbb{R}^d)$ is said to be **Cauchy** if $d(f_n, f_m) \to 0$ as $n, m \to \infty$. Moreover, this sequence converges to $f \in L^2(\mathbb{R}^d)$ if $d(f_n, f) \to 0$ as $n \to \infty$.

Theorem 1.2 *The space $L^2(\mathbb{R}^d)$ is complete in its metric.*

In other words, every Cauchy sequence in $L^2(\mathbb{R}^d)$ converges to a function in $L^2(\mathbb{R}^d)$. This theorem, which is in sharp contrast with the situation for Riemann integrable functions, is a graphic illustration of the usefulness of Lebesgue's theory of integration. We elaborate on this point and its relation to Fourier series in Section 3 below.

Proof. The argument given here follows closely the proof in Chapter 2 that L^1 is complete. Let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence in L^2, and consider a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ of $\{f_n\}$ with the following property:

$$\|f_{n_{k+1}} - f_{n_k}\| \leq 2^{-k}, \qquad \text{for all } k \geq 1.$$

If we now consider the series whose convergence will be seen below,

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty} |(f_{n_{k+1}}(x) - f_{n_k}(x))|,$$

together the partial sums

$$S_K(f)(x) = f_{n_1}(x) + \sum_{k=1}^{K} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$S_K(g)(x) = |f_{n_1}(x)| + \sum_{k=1}^{K} |f_{n_{k+1}}(x) - f_{n_k}(x)|,$$

then the triangle inequality implies

$$\begin{aligned}\|S_K(g)\| &\leq \|f_{n_1}\| + \sum_{k=1}^{K} \|f_{n_{k+1}} - f_{n_k}\| \\ &\leq \|f_{n_1}\| + \sum_{k=1}^{K} 2^{-k}.\end{aligned}$$

Letting K tend to infinity, and applying the monotone convergence theorem proves that $\int |g|^2 < \infty$, and since $|f| \leq g$, we must have $f \in L^2(\mathbb{R}^d)$.

In particular, the series defining f converges almost everywhere, and since (by construction of the telescopic series) the $(K-1)^{\text{th}}$ partial sum of this series is precisely f_{n_K}, we find that

$$f_{n_k}(x) \to f(x) \qquad \text{a.e. } x.$$

To prove that $f_{n_k} \to f$ in $L^2(\mathbb{R}^d)$ as well, we simply observe that $|f - S_K(f)|^2 \leq (2g)^2$ for all K, and apply the dominated convergence theorem to get $\|f_{n_k} - f\| \to 0$ as k tends to infinity.

Finally, the last step of the proof consists of recalling that $\{f_n\}$ is Cauchy. Given ϵ, there exists N such that for all $n, m > N$ we have $\|f_n - f_m\| < \epsilon/2$. If n_k is chosen so that $n_k > N$, and $\|f_{n_k} - f\| < \epsilon/2$, then the triangle inequality implies

$$\|f_n - f\| \leq \|f_n - f_{n_k}\| + \|f_{n_k} - f\| < \epsilon$$

whenever $n > N$. This concludes the proof of the theorem.

An additional useful property of $L^2(\mathbb{R}^d)$ is contained in the following theorem.

Theorem 1.3 *The space $L^2(\mathbb{R}^d)$ is **separable**, in the sense that there exists a countable collection $\{f_k\}$ of elements in $L^2(\mathbb{R}^d)$ such that their linear combinations are dense in $L^2(\mathbb{R}^d)$.*

Proof. Consider the family of functions of the form $r\chi_R(x)$, where r is a complex number with rational real and imaginary parts, and R is a rectangle in $\mathbb{R}^d$ with rational coordinates. We claim that finite linear combinations of these type of functions are dense in $L^2(\mathbb{R}^d)$.

Suppose $f \in L^2(\mathbb{R}^d)$ and let $\epsilon > 0$. Consider for each $n \geq 1$ the function g_n defined by

$$g_n(x) = \begin{cases} f(x) & \text{if } |x| \leq n \text{ and } |f(x)| \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $|f - g_n|^2 \leq 4|f|^2$ and $g_n(x) \to f(x)$ almost everywhere.[1] The dominated convergence theorem implies that $\|f - g_n\|^2_{L^2(\mathbb{R}^d)} \to 0$ as n tends to infinity; therefore we have

$$\|f - g_N\|_{L^2(\mathbb{R}^d)} < \epsilon/2 \quad \text{for some } N.$$

Let $g = g_N$, and note that g is a bounded function supported on a bounded set; thus $g \in L^1(\mathbb{R}^d)$. We may now find a step function φ so that $|\varphi| \leq N$ and $\int |g - \varphi| < \epsilon^2/16N$ (Theorem 2.4, Chapter 2). By replacing the coefficients and rectangles that appear in the canonical form of φ by complex numbers with rational real and imaginary parts, and rectangles with rational coordinates, we may find a ψ with $|\psi| \leq N$ and $\int |g - \psi| < \epsilon^2/8N$. Finally, we note that

$$\int |g - \psi|^2 \leq 2N \int |g - \psi| < \epsilon^2/4.$$

Consequently $\|g - \psi\| < \epsilon/2$, therefore $\|f - \psi\| < \epsilon$, and the proof is complete.

The example $L^2(\mathbb{R}^d)$ possesses all the characteristic properties of a Hilbert space, and motivates the definition of the abstract version of this concept.

2 Hilbert spaces

A set $\mathcal{H}$ is a **Hilbert space** if it satisfies the following:

(i) $\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$).[2]

(ii) $\mathcal{H}$ is equipped with an inner product $(\cdot, \cdot)$, so that

- $f \mapsto (f, g)$ is linear on $\mathcal{H}$ for every fixed $g \in \mathcal{H}$,
- $(f, g) = \overline{(g, f)}$,
- $(f, f) \geq 0$ for all $f \in \mathcal{H}$.

We let $\|f\| = (f, f)^{1/2}$.

(iii) $\|f\| = 0$ if and only if $f = 0$.

[1] By definition $f \in L^2(\mathbb{R}^d)$ implies that $|f|^2$ is integrable, hence $f(x)$ is finite for a.e x.

[2] At this stage we consider both cases, where the scalar field can be either $\mathbb{C}$ or $\mathbb{R}$. However, in many applications, such as in the context of Fourier analysis, one deals primarily with Hilbert spaces over $\mathbb{C}$.

(iv) The Cauchy-Schwarz and triangle inequalities hold

$$|(f, g)| \leq \|f\| \, \|g\| \quad \text{and} \quad \|f + g\| \leq \|f\| + \|g\|$$

for all $f, g \in \mathcal{H}$.

(v) $\mathcal{H}$ is complete in the metric $d(f, g) = \|f - g\|$.

(vi) $\mathcal{H}$ is separable.

We make two comments about the definition of a Hilbert space. First, the Cauchy-Schwarz and triangle inequalities in (iv) are in fact easy consequences of assumptions (i) and (ii). (See Exercise 1.) Second, we make the requirement that $\mathcal{H}$ be separable because that is the case in most applications encountered. That is not to say that there are no interesting non-separable examples; one such example is described in Problem 2.

Also, we remark that in the context of a Hilbert space we shall often write $\lim_{n\to\infty} f_n = f$ or $f_n \to f$ to mean that $\lim_{n\to\infty} \|f_n - f\| = 0$, which is the same as $d(f_n, f) \to 0$.

We give some examples of Hilbert spaces.

Example 1. If E is a measurable subset of $\mathbb{R}^d$ with $m(E) > 0$, we let $L^2(E)$ denote the space of square integrable functions that are supported on E,

$$L^2(E) = \left\{ f \text{ supported on } E, \text{ so that } \int_E |f(x)|^2 \, dx < \infty \right\}.$$

The inner product and norm on $L^2(E)$ are then

$$(f, g) = \int_E f(x)\overline{g(x)} \, dx \quad \text{and} \quad \|f\| = \left(\int_E |f(x)|^2 \, dx \right)^{1/2}.$$

Once again, we consider two elements of $L^2(E)$ to be equivalent if they differ only on a set of measure zero; this guarantees that $\|f\| = 0$ implies $f = 0$. The properties (i) through (vi) follow from these of $L^2(\mathbb{R}^d)$ proved above.

Example 2. A simple example is the finite-dimensional complex Euclidean space. Indeed,

$$\mathbb{C}^N = \{(a_1, \ldots, a_N) : \; a_k \in \mathbb{C}\}$$

becomes a Hilbert space when equipped with the inner product

$$\sum_{k=1}^{N} a_k \overline{b_k},$$

where $a = (a_1, \ldots, a_N)$ and $b = (b_1, \ldots, b_N)$ are in $\mathbb{C}^N$. The norm is then

$$\|a\| = \left(\sum_{k=1}^{N} |a_k|^2 \right)^{1/2}.$$

One can formulate in the same way the real Hilbert space $\mathbb{R}^N$.

Example 3. An infinite-dimensional analogue of the above example is the space $\ell^2(\mathbb{Z})$. By definition

$$\ell^2(\mathbb{Z}) = \left\{ (\ldots, a_{-2}, a_{-1}, a_0, a_1, \ldots) : \ a_i \in \mathbb{C}, \quad \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty \right\}.$$

If we denote infinite sequences by a and b, the inner product and norm on $\ell^2(\mathbb{Z})$ are

$$(a, b) = \sum_{k=-\infty}^{\infty} a_k \overline{b_k} \quad \text{and} \quad \|a\| = \left(\sum_{k=-\infty}^{\infty} |a_k|^2 \right)^{1/2}.$$

We leave the proof that $\ell^2(\mathbb{Z})$ is a Hilbert space as Exercise 4.

While this example is very simple, it will turn out that all infinite-dimensional (separable) Hilbert spaces are $\ell^2(\mathbb{Z})$ in disguise.

Also, a slight variant of this space is $\ell^2(\mathbb{N})$, where we take only one-sided sequences, that is,

$$\ell^2(\mathbb{N}) = \left\{ (a_1, a_2, \ldots) : \ a_i \in \mathbb{C}, \ \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

The inner product and norm are then defined in the same way with the sums extending from $n = 1$ to ∞.

A characteristic feature of a Hilbert space is the notion of orthogonality. This aspect, with its rich geometric and analytic consequences, distinguishes Hilbert spaces from other normed vector spaces. We now describe some of these properties.

2.1 Orthogonality

Two elements f and g in a Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)$ are **orthogonal** or **perpendicular** if

$$(f,g)=0, \qquad \text{and we then write } f \perp g.$$

The first simple observation is that the usual theorem of Pythagoras holds in the setting of abstract Hilbert spaces:

Proposition 2.1 *If $f \perp g$, then $\|f+g\|^2 = \|f\|^2 + \|g\|^2$.*

Proof. It suffices to note that $(f,g)=0$ implies $(g,f)=0$, and therefore

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

A finite or countably infinite subset $\{e_1, e_2, \ldots\}$ of a Hilbert space $\mathcal{H}$ is **orthonormal** if

$$(e_k, e_\ell) = \begin{cases} 1 & \text{when } k=\ell, \\ 0 & \text{when } k \neq \ell. \end{cases}$$

In other words, each e_k has unit norm and is orthogonal to e_ℓ whenever $\ell \neq k$.

Proposition 2.2 *If $\{e_k\}_{k=1}^\infty$ is orthonormal, and $f = \sum a_k e_k \in \mathcal{H}$ where the sum is finite, then*

$$\|f\|^2 = \sum |a_k|^2.$$

The proof is a simple application of the Pythagorean theorem.

Given an orthonormal subset $\{e_1, e_2, \ldots\} = \{e_k\}_{k=1}^\infty$ of $\mathcal{H}$, a natural problem is to determine whether this subset spans all of $\mathcal{H}$, that is, whether finite linear combinations of elements in $\{e_1, e_2, \ldots\}$ are dense in $\mathcal{H}$. If this is the case, we say that $\{e_k\}_{k=1}^\infty$ is an **orthonormal basis** for $\mathcal{H}$. If we are in the presence of an orthonormal basis, we might expect that any $f \in \mathcal{H}$ takes the form

$$f = \sum_{k=1}^\infty a_k e_k,$$

for some constants $a_k \in \mathbb{C}$. In fact, taking the inner product of both sides with e_j, and recalling that $\{e_k\}$ is orthonormal yields (formally)

$$(f, e_j) = a_j.$$

This question is motivated by Fourier series. In fact, a good insight into the theorem below is afforded by considering the case where $\mathcal{H}$ is $L^2([-\pi,\pi])$ with inner product $(f,g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$, and the orthonormal set $\{e_k\}_{k=1}^{\infty}$ is merely a relabeling of the exponentials $\{e^{inx}\}_{n=-\infty}^{\infty}$.

Adapting the notation used in Fourier series, we write $f \sim \sum_{k=1}^{\infty} a_k e_k$, where $a_j = (f, e_j)$ for all j.

In the next theorem, we provide four equivalent characterizations that $\{e_k\}$ is an orthonormal basis for $\mathcal{H}$.

Theorem 2.3 *The following properties of an orthonormal set* $\{e_k\}_{k=1}^{\infty}$ *are equivalent.*

(i) *Finite linear combinations of elements in* $\{e_k\}$ *are dense in* $\mathcal{H}$.

(ii) *If* $f \in \mathcal{H}$ *and* $(f, e_j) = 0$ *for all* j*, then* $f = 0$.

(iii) *If* $f \in \mathcal{H}$*, and* $S_N(f) = \sum_{k=1}^{N} a_k e_k$*, where* $a_k = (f, e_k)$*, then* $S_N(f) \to f$ *as* $N \to \infty$ *in the norm.*

(iv) *If* $a_k = (f, e_k)$*, then* $\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2$.

Proof. We prove that each property implies the next, with the last one implying the first.

We begin by assuming (i). Given $f \in \mathcal{H}$ with $(f, e_j) = 0$ for all j, we wish to prove that $f = 0$. By assumption, there exists a sequence $\{g_n\}$ of elements in $\mathcal{H}$ that are finite linear combinations of elements in $\{e_k\}$, and such that $\|f - g_n\|$ tends to 0 as n goes to infinity. Since $(f, e_j) = 0$ for all j, we must have $(f, g_n) = 0$ for all n; therefore an application of the Cauchy-Schwarz inequality gives

$$\|f\|^2 = (f,f) = (f, f - g_n) \le \|f\|\,\|f - g_n\| \quad \text{for all } n.$$

Letting $n \to \infty$ proves that $\|f\|^2 = 0$; hence $f = 0$, and (i) implies (ii).

Now suppose that (ii) is verified. For $f \in \mathcal{H}$ we define

$$S_N(f) = \sum_{k=1}^{N} a_k e_k, \qquad \text{where } a_k = (f, e_k),$$

and prove first that $S_N(f)$ converges to some element $g \in \mathcal{H}$. Indeed, one notices that the definition of a_k implies $(f - S_N(f)) \perp S_N(f)$, so the Pythagorean theorem and Proposition 2.2 give

$$\|f\|^2 = \|f - S_N(f)\|^2 + \|S_N(f)\|^2 = \|f - S_N(f)\|^2 + \sum_{k=1}^{N} |a_k|^2. \tag{2}$$

Hence $\|f\|^2 \geq \sum_{k=1}^N |a_k|^2$, and letting N tend to infinity we obtain **Bessel's inequality**

$$\sum_{k=1}^{\infty} |a_k|^2 \leq \|f\|^2,$$

which implies that the series $\sum_{k=1}^{\infty} |a_k|^2$ converges. Therefore, $\{S_N(f)\}_{N=1}^{\infty}$ forms a Cauchy sequence in $\mathcal{H}$ since

$$\|S_N(f) - S_M(f)\|^2 = \sum_{k=M+1}^{N} |a_k|^2 \qquad \text{whenever } N > M.$$

Since $\mathcal{H}$ is complete, there exists $g \in \mathcal{H}$ such that $S_N(f) \to g$ as N tends to infinity.

Fix j, and note that for all sufficiently large N, $(f - S_N(f), e_j) = a_j - a_j = 0$. Since $S_N(f)$ tends to g, we conclude that

$$(f - g, e_j) = 0 \qquad \text{for all } j.$$

Hence $f = g$ by assumption (ii), and we have proved that $f = \sum_{k=1}^{\infty} a_k e_k$.

Now assume that (iii) holds. Observe from (2) that we immediately get in the limit as N goes to infinity

$$\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2.$$

Finally, if (iv) holds, then again from (2) we see that $\|f - S_N(f)\|$ converges to 0. Since each $S_N(f)$ is a finite linear combination of elements in $\{e_k\}$, we have completed the circle of implications, and the theorem is proved.

In particular, a closer look at the proof shows that Bessel's inequality holds for any orthonormal family $\{e_k\}$. In contrast, the identity

$$\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2, \qquad \text{where } a_k = (f, e_k),$$

which is called **Parseval's identity**, holds if and only if $\{e_k\}_{k=1}^{\infty}$ is also an orthonormal basis.

Now we turn our attention to the existence of a basis.

Theorem 2.4 *Any Hilbert space has an orthonormal basis.*

The first step in the proof of this fact is to recall that (by definition) a Hilbert space $\mathcal{H}$ is separable. Hence, we may choose a countable collection of elements $\mathcal{F} = \{h_k\}$ in $\mathcal{H}$ so that finite linear combinations of elements in $\mathcal{F}$ are dense in $\mathcal{H}$.

We start by recalling a definition already used in the case of finite-dimensional vector spaces. Finitely many elements $g_1, \ldots, g_N$ are said to be **linearly independent** if whenever

$$a_1 g_1 + \cdots + a_N g_N = 0 \qquad \text{for some complex numbers } a_i,$$

then $a_1 = a_2 = \cdots = a_N = 0$. In other words, no element g_j is a linear combination of the others. In particular, we note that none of the g_j can be 0. We say that a countable family of elements is **linearly independent** if all finite subsets of this family are linearly independent.

If we next successively disregard the elements h_k that are linearly dependent on the previous elements $h_1, h_2, \ldots, h_{k-1}$, then the resulting collection $h_1 = f_1, f_2, \ldots, f_k, \ldots$ consists of linearly independent elements, whose finite linear combinations are the same as those given by $h_1, h_2, \ldots, h_k, \ldots$, and hence these linear combinations are also dense in $\mathcal{H}$.

The proof of the theorem now follows from an application of a familiar construction called the **Gram-Schmidt process**. Given a finite family of elements $\{f_1, \ldots, f_k\}$ we call the **span** of this family the set of all elements which are finite linear combinations of the elements $\{f_1, \ldots, f_k\}$. We denote the span of $\{f_1, \ldots, f_k\}$ by $\mathrm{Span}(\{f_1, \ldots, f_k\})$.

We now construct a sequence of orthonormal vectors $e_1, e_2, \ldots$ such that $\mathrm{Span}(\{e_1, \ldots, e_n\}) = \mathrm{Span}(\{f_1, \ldots, f_n\})$ for all $n \geq 1$. We do this by induction.

By the linear independence hypothesis, $f_1 \neq 0$, so we may take $e_1 = f_1/\|f_1\|$. Next, assume that orthonormal vectors $e_1, \ldots, e_k$ have been found such that $\mathrm{Span}(\{e_1, \ldots, e_k\}) = \mathrm{Span}(\{f_1, \ldots, f_k\})$ for a given k. We then try e'_{k+1} as $f_{k+1} + \sum_{j=1}^{k} a_j e_j$. To have $(e'_{k+1}, e_j) = 0$ requires that $a_j = -(f_{k+1}, e_j)$, and this choice of a_j for $1 \leq j \leq k$ assures that e'_{k+1} is orthogonal to $e_1, \ldots, e_k$. Moreover our linear independence hypothesis assures that $e'_{k+1} \neq 0$; hence we need only "renormalize" and take $e_{k+1} = e'_{k+1}/\|e'_{k+1}\|$ to complete the inductive step. With this we have found an orthonormal basis for $\mathcal{H}$

Note that we have implicitly assumed that the number of linearly independent elements $f_1, f_2, \ldots$ is infinite. In the case where there are only N linearly independent vectors $f_1, \ldots, f_N$, then $e_1, \ldots, e_N$ constructed in the same way also provide an orthonormal basis for $\mathcal{H}$. These two cases are differentiated in the following definition. If $\mathcal{H}$ is a Hilbert space with an orthonormal basis consisting of finitely many elements, then we say that $\mathcal{H}$ is **finite-dimensional**. Otherwise $\mathcal{H}$ is said to be **infinite-dimensional**.

2.2 Unitary mappings

A correspondence between two Hilbert spaces that preserves their structure is a unitary transformation. More precisely, suppose we are given two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ with respective inner products $(\cdot,\cdot)_{\mathcal{H}}$ and $(\cdot,\cdot)_{\mathcal{H}'}$, and the corresponding norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}'}$. A mapping $U : \mathcal{H} \to \mathcal{H}'$ between these space is called **unitary** if:

(i) U is linear, that is, $U(\alpha f + \beta g) = \alpha U(f) + \beta U(g)$.

(ii) U is a bijection.

(iii) $\|Uf\|_{\mathcal{H}'} = \|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$.

Some observations are in order. First, since U is bijective it must have an inverse $U^{-1} : \mathcal{H}' \to \mathcal{H}$ that is also unitary. Part (iii) above also implies that if U is unitary, then

$$(Uf, Ug)_{\mathcal{H}'} = (f,g)_{\mathcal{H}} \qquad \text{for all } f, g \in \mathcal{H}.$$

To see this, it suffices to "polarize," that is, to note that for any vector space (say over $\mathbb{C}$) with inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$, we have

$$(F,G) = \frac{1}{4}\left[\|F+G\|^2 - \|F-G\|^2 + i\left(\|\frac{F}{i} + G\|^2 - \|\frac{F}{i} - G\|^2\right)\right]$$

whenever F and G are elements of the space.

The above leads us to say that the two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ are **unitarily equivalent** or **unitarily isomorphic** if there exists a unitary mapping $U : \mathcal{H} \to \mathcal{H}'$. Clearly, unitary isomorphism of Hilbert spaces is an equivalence relation.

With this definition we are now in a position to give precise meaning to the statement we made earlier that all infinite-dimensional Hilbert spaces are the same and in that sense $\ell^2(\mathbb{Z})$ in disguise.

Corollary 2.5 *Any two infinite-dimensional Hilbert spaces are unitarily equivalent.*

Proof. If $\mathcal{H}$ and $\mathcal{H}'$ are two infinite-dimensional Hilbert spaces, we may select for each an orthonormal basis, say

$$\{e_1, e_2, \ldots\} \subset \mathcal{H} \quad \text{and} \quad \{e_1', e_2', \ldots\} \subset \mathcal{H}'.$$

Then, consider the mapping defined as follows: if $f = \sum_{k=1}^{\infty} a_k e_k$, then

$$U(f) = g, \quad \text{where} \quad g = \sum_{k=1}^{\infty} a_k e_k'.$$

Clearly, the mapping U is both linear and invertible. Moreover, by Parseval's identity, we must have

$$\|Uf\|_{\mathcal{H}'}^2 = \|g\|_{\mathcal{H}'}^2 = \sum_{k=1}^{\infty} |a_k|^2 = \|f\|_{\mathcal{H}}^2,$$

and the corollary is proved.

Consequently, all infinite-dimensional Hilbert spaces are unitarily equivalent to $\ell^2(\mathbb{N})$, and thus, by relabeling, to $\ell^2(\mathbb{Z})$. By similar reasoning we also have the following:

Corollary 2.6 *Any two finite-dimensional Hilbert spaces are unitarily equivalent if and only if they have the same dimension.*

Thus every finite-dimensional Hilbert space over $\mathbb{C}$ (or over $\mathbb{R}$) is equivalent with $\mathbb{C}^d$ (or $\mathbb{R}^d$), for some d.

2.3 Pre-Hilbert spaces

Although Hilbert spaces arise naturally, one often starts with a **pre-Hilbert space** instead, that is, a space $\mathcal{H}_0$ that satisfies all the defining properties of a Hilbert space except (v); in other words $\mathcal{H}_0$ is not assumed to be complete. A prime example arose implicitly early in the study of Fourier series with the space $\mathcal{H}_0 = \mathcal{R}$ of Riemann integrable functions on $[-\pi, \pi]$ with the usual inner product; we return to this below. Other examples appear in the next chapter in the study of the solutions of partial differential equations.

Fortunately, every pre-Hilbert space $\mathcal{H}_0$ can be completed.

Proposition 2.7 *Suppose we are given a pre-Hilbert space $\mathcal{H}_0$ with inner product $(\cdot,\cdot)_0$. Then we can find a Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)$ such that*

(i) $\mathcal{H}_0 \subset \mathcal{H}$.

(ii) $(f,g)_0 = (f,g)$ *whenever* $f,g \in \mathcal{H}_0$.

(iii) $\mathcal{H}_0$ *is dense in* $\mathcal{H}$.

A Hilbert space satisfying properties like $\mathcal{H}$ in the above proposition is called a **completion** of $\mathcal{H}_0$. We shall only sketch the construction of $\mathcal{H}$, since it follows closely Cantor's familiar method of obtaining the real numbers as the completion of the rationals in terms of Cauchy sequences of rationals.

Indeed, consider the collection of all Cauchy sequences $\{f_n\}$ with $f_n \in \mathcal{H}_0$, $1 \leq n < \infty$. One defines an equivalence relation in this collection by saying that $\{f_n\}$ is equivalent to $\{f'_n\}$ if $f_n - f'_n$ converges to 0 as $n \to \infty$. The collection of equivalence classes is then taken to be $\mathcal{H}$. One then easily verifies that $\mathcal{H}$ inherits the structure of a vector space, with an inner product (f,g) defined as $\lim_{n\to\infty}(f_n, g_n)$, where $\{f_n\}$ and $\{g_n\}$ are Cauchy sequences in $\mathcal{H}_0$, representing, respectively, the elements f and g in $\mathcal{H}$. Next, if $f \in \mathcal{H}_0$ we take the sequence $\{f_n\}$, with $f_n = f$ for all n, to represent f as an element of $\mathcal{H}$, giving $\mathcal{H}_0 \subset \mathcal{H}$. To see that $\mathcal{H}$ is complete, let $\{F^k\}_{k=1}^{\infty}$ be a Cauchy sequence in $\mathcal{H}$, with each F^k represented by $\{f_n^k\}_{n=1}^{\infty}$, $f_n^k \in \mathcal{H}_0$. If we define $F \in \mathcal{H}$ as represented by the sequence $\{f_n\}$ with $f_n = f^n_{N(n)}$, where $N(n)$ is so that $|f^n_{N(n)} - f^n_j| \leq 1/n$ for $j \geq N(n)$, then we note that $F^k \to F$ in $\mathcal{H}$.

One can also observe that the completion $\mathcal{H}$ of $\mathcal{H}_0$ is unique up to isomorphism. (See Exercise 14.)

3 Fourier series and Fatou's theorem

We have already seen an interesting relation between Hilbert spaces and some elementary facts about Fourier series. Here we want to pursue this idea and also connect it with complex analysis.

When considering Fourier series, it is natural to begin by turning to the broader class of all integrable functions on $[-\pi,\pi]$. Indeed, note that $L^2([-\pi,\pi]) \subset L^1([-\pi,\pi])$, by the Cauchy-Schwarz inequality, since the interval $[-\pi,\pi]$ has finite measure. Thus, if $f \in L^1([-\pi,\pi])$ and $n \in \mathbb{Z}$, we define the n^{th} **Fourier coefficient** of f by

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx.$$

The **Fourier series** of f is then formally $\sum_{n=-\infty}^{\infty} a_n e^{inx}$, and we write

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}$$

to indicate that the sum on the right is the Fourier series of the function on the left. The theory developed thus far provides the natural generalization of some earlier results obtained in Book I.

Theorem 3.1 *Suppose f is integrable on $[-\pi, \pi]$.*

(i) *If $a_n = 0$ for all n, then $f(x) = 0$ for a.e. x.*

(ii) *$\sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{inx}$ tends to $f(x)$ for a.e. x, as $r \to 1$, $r < 1$.*

The second conclusion is the almost everywhere "Abel summability" to f of its Fourier series. Note that since $|a_n| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|\,dx$, the series $\sum a_n r^{|n|} e^{inx}$ converges absolutely and uniformly for each r, $0 \leq r < 1$.

Proof. The first conclusion is an immediate consequence of the second. To prove the latter we recall the identity

$$\sum_{n=-\infty}^{\infty} r^{|n|} e^{iny} = P_r(y) = \frac{1 - r^2}{1 - 2r\cos y + r^2}$$

for the Poisson kernel; see Book I, Chapter 2. Starting with our given $f \in L^1([-\pi, \pi])$ we extend it as a function on $\mathbb{R}$ by making it periodic of period 2π.[3] We then claim that for every x

$$\sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{inx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y) P_r(y)\,dy. \tag{3}$$

Indeed, by the dominated convergence theorem the right-hand side equals

$$\sum r^{|n|} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y) e^{iny}\,dy.$$

Moreover, for each x and n

$$\begin{aligned}\int_{-\pi}^{\pi} f(x-y)e^{iny}\,dy &= \int_{-\pi+x}^{\pi+x} f(y)e^{in(x-y)}\,dy \\ &= e^{inx}\int_{-\pi}^{\pi} f(y)e^{-iny}\,dy = e^{inx} 2\pi a_n.\end{aligned}$$

[3] Note that we may without loss of generality assume that $f(\pi) = f(-\pi)$ so as to make the periodic extension unambiguous.

The first equality follows by translation invariance (see Section 3, Chapter 2), and the second since $\int_{-\pi}^{\pi} F(y)\,dy = \int_I F(y)\,dy$ whenever F is periodic of period 2π and I is an interval of length 2π (Exercise 3, Chapter 2). With these observations, the identity (3) is established. We can now invoke the facts about approximations to the identity (Theorem 2.1 and Example 4, Chapter 3) to conclude that the left-hand side of (3) tends to $f(x)$ at every point of the Lebesgue set of f, hence almost everywhere. (To be correct, the hypotheses of the theorem require that f be integrable on all of $\mathbb{R}$. We can achieve this for our periodic function by setting f equal to zero outside $[-2\pi, 2\pi]$, and then (3) still holds for this modified f, whenever $x \in [-\pi, \pi]$.)

We return to the more restrictive setting of L^2. We express the essential conclusions of Theorem 2.3 in the context of Fourier series. With $f \in L^2([-\pi, \pi])$, we write as before $a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$.

Theorem 3.2 *Suppose $f \in L^2([-\pi, \pi])$. Then:*

(i) *We have Parseval's relation*

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx.$$

(ii) *The mapping $f \mapsto \{a_n\}$ is a unitary correspondence between $L^2([-\pi, \pi])$ and $\ell^2(\mathbb{Z})$.*

(iii) *The Fourier series of f converges to f in the L^2-norm, that is,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - S_N(f)(x)|^2\,dx \to 0 \quad \text{as } N \to \infty,$$

where $S_N(f) = \sum_{|n|\leq N} a_n e^{inx}$.

To apply the previous results, we let $\mathcal{H} = L^2([-\pi, \pi])$ with inner product $(f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$, and take the orthonormal set $\{e_k\}_{k=1}^{\infty}$ to be the exponentials $\{e^{inx}\}_{n=-\infty}^{\infty}$, with $k = 1$ when $n = 0$, $k = 2n$ for $n > 0$, and $k = 2|n| - 1$ for $n < 0$.

By the previous result, assertion (ii) of Theorem 2.3 holds and thus all the other conclusions hold. We therefore have Parseval's relation, and from (iv) we conclude that $\|f - S_N(f)\|^2 = \sum_{|n|>N} |a_n|^2 \to 0$ as $N \to \infty$. Similarly, if $\{a_n\} \in \ell^2(\mathbb{Z})$ is given, then $\|S_N(f) - S_M(f)\|^2 \to 0$, as $N, M \to \infty$. Hence the completeness of L^2 guarantees that there is an $f \in L^2$ such that $\|f - S_N(f)\| \to 0$, and one verifies directly that f

has $\{a_n\}$ as its Fourier coefficients. Thus we deduce that the mapping $f \mapsto \{a_n\}$ is onto and hence unitary. This is a key conclusion that holds in the setting on L^2 and was not valid in an earlier context of Riemann integrable functions. In fact the space $\mathcal{R}$ of such functions on $[-\pi, \pi]$ is not complete in the norm, containing as it does the continuous functions, but $\mathcal{R}$ is itself restricted to bounded functions.

3.1 Fatou's theorem

Fatou's theorem is a remarkable result in complex analysis. Its proof combines elements of Hilbert spaces, Fourier series, and deeper ideas of differentiation theory, and yet none of these notions appear in its statement. The question that Fatou's theorem answers may be put simply as follows.

> Suppose $F(z)$ is holomorphic in the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. What are conditions on F that guarantee that $F(z)$ will converge, in an appropriate sense, to boundary values $F(e^{i\theta})$ on the unit circle?

In general a holomorphic function in the unit disc can behave quite erratically near the boundary. It turns out, however, that imposing a simple boundedness condition is enough to obtain a strong conclusion.

If F is a function defined in the unit disc $\mathbb{D}$, we say that F has a **radial limit** at the point $-\pi \leq \theta \leq \pi$ on the circle, if the limit

$$\lim_{\substack{r \to 1 \\ r < 1}} F(re^{i\theta})$$

exists.

Theorem 3.3 *A bounded holomorphic function $F(re^{i\theta})$ on the unit disc has radial limits at almost every θ.*

Proof. We know that $F(z)$ has a power series expansion $\sum_{n=0}^{\infty} a_n z^n$ in $\mathbb{D}$ that converges absolutely and uniformly whenever $z = re^{i\theta}$ and $r < 1$. In fact, for $r < 1$ the series $\sum_{n=0}^{\infty} a_n r^n e^{in\theta}$ is the Fourier series of the function $F(re^{i\theta})$, that is,

$$a_n r^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(re^{i\theta}) e^{-in\theta} \, d\theta \qquad \text{when } n \geq 0,$$

and the integral vanishes when $n < 0$. (See also Chapter 3, Section 7 in Book II).

We pick M so that $|F(z)| \le M$, for all $z \in \mathbb{D}$. By Parseval's identity

$$\sum_{n=0}^{\infty} |a_n|^2 r^{2n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |F(re^{i\theta})|^2 d\theta \qquad \text{for each } 0 \le r < 1.$$

Letting $r \to 1$ one sees that $\sum |a_n|^2$ converges (and is $\le M^2$). We now let $F(e^{i\theta})$ be the L^2-function whose Fourier coefficients are a_n when $n \ge 0$, and 0 when $n < 0$. Hence by conclusion (ii) in Theorem 3.1

$$\sum_{n=0}^{\infty} a_n r^n e^{in\theta} \to F(e^{i\theta}), \qquad \text{for a.e } \theta,$$

concluding the proof of the theorem.

If we examine the argument given above we see that the same conclusion holds for a larger class of functions. In this connection, we define the **Hardy space** $H^2(\mathbb{D})$ to consist of all holomorphic functions F on the unit disc $\mathbb{D}$ that satisfy

$$\sup_{0\le r<1} \frac{1}{2\pi}\int_{-\pi}^{\pi} |F(re^{i\theta})|^2 \, d\theta < \infty.$$

We also define the "norm" for functions F in this class, $\|F\|_{H^2(\mathbb{D})}$, to be the square root of the above quantity.

One notes that if F is bounded, then $F \in H^2(\mathbb{D})$, and moreover the conclusion of the existence of radial limits almost everywhere holds for any $F \in H^2(\mathbb{D})$, by the same argument given for the bounded case.[4] Finally, one notes that $F \in H^2(\mathbb{D})$ if and only if $F(z) = \sum_{n=0}^{\infty} a_n z^n$ with $\sum_{n=0}^{\infty} |a_n|^2 < \infty$; moreover, $\sum_{n=0}^{\infty} |a_n|^2 = \|F\|^2_{H^2(\mathbb{D})}$. This states in particular that $H^2(\mathbb{D})$ is in fact a Hilbert space that can be viewed as the "subspace" $\ell^2(\mathbb{Z}^+)$ of $\ell^2(\mathbb{Z})$, consisting of all $\{a_n\} \in \ell^2(\mathbb{Z})$, with $a_n = 0$ when $n < 0$.

Some general considerations of subspaces and their concomitant orthogonal projections will be taken up next.

4 Closed subspaces and orthogonal projections

A **linear subspace** $\mathcal{S}$ (or simply subspace) of $\mathcal{H}$ is a subset of $\mathcal{H}$ that satisfies $\alpha f + \beta g \in \mathcal{S}$ whenever $f, g \in \mathcal{S}$ and α, β are scalars. In other words, $\mathcal{S}$ is also a vector space. For example in $\mathbb{R}^3$, lines passing through

[4] An even more general statement is given in Problem 5*.

the origin and planes passing through the origin are the one-dimensional and two-dimensional subspaces, respectively.

The subspace $\mathcal{S}$ is **closed** if whenever $\{f_n\} \subset \mathcal{S}$ converges to some $f \in \mathcal{H}$, then f also belongs to $\mathcal{S}$. In the case of finite-dimensional Hilbert spaces, every subspace is closed. This is, however, not true in the general case of infinite-dimensional Hilbert spaces. For instance, as we have already indicated, the subspace of Riemann integrable functions in $L^2([-\pi, \pi])$ is not closed, nor is the subspace obtained by fixing a basis and taking all vectors that are finite linear combinations of these basis elements. It is useful to note that every closed subspace $\mathcal{S}$ of $\mathcal{H}$ is itself a Hilbert space, with the inner product on $\mathcal{S}$ that which is inherited from $\mathcal{H}$. (For the separability of $\mathcal{S}$, see Exercise 11.)

Next, we show that a closed subspace enjoys an important characteristic property of Euclidean geometry.

Lemma 4.1 *Suppose $\mathcal{S}$ is a closed subspace of $\mathcal{H}$ and $f \in \mathcal{H}$. Then:*

(i) *There exists a (unique) element $g_0 \in \mathcal{S}$ which is closest to f, in the sense that*

$$\|f - g_0\| = \inf_{g \in \mathcal{S}} \|f - g\|.$$

(ii) *The element $f - g_0$ is perpendicular to $\mathcal{S}$, that is,*

$$(f - g_0, g) = 0 \quad \textit{for all } g \in \mathcal{S}.$$

The situation in the lemma can be visualized as in Figure 1.

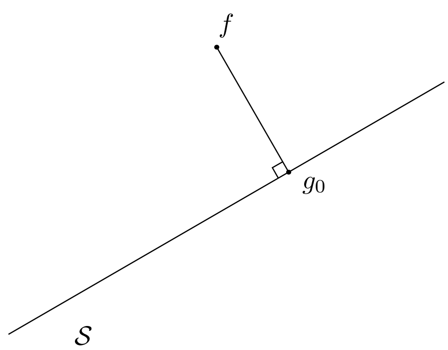

Figure 1. Nearest element to f in $\mathcal{S}$

Proof. If $f \in \mathcal{S}$, then we choose $f = g_0$, and there is nothing left to prove. Otherwise, we let $d = \inf_{g\in\mathcal{S}} \|f - g\|$, and note that we must have $d > 0$ since $f \notin \mathcal{S}$ and $\mathcal{S}$ is closed. Consider a sequence $\{g_n\}_{n=1}^{\infty}$ in $\mathcal{S}$ such that

$$\|f - g_n\| \to d \qquad \text{as } n \to \infty.$$

We claim that $\{g_n\}$ is a Cauchy sequence whose limit will be the desired element g_0. In fact, it would suffice to show that a subsequence of $\{g_n\}$ converges, and this is immediate in the finite-dimensional case because a closed ball is compact. However, in general this compactness fails, as we shall see in Section 6, and so a more intricate argument is needed at this point.

To prove our claim, we use the **parallelogram law**, which states that in a Hilbert space $\mathcal{H}$

$$\|A + B\|^2 + \|A - B\|^2 = 2\left[\|A\|^2 + \|B\|^2\right] \qquad \text{for all } A, B \in \mathcal{H}. \tag{4}$$

The simple verification of this equality, which consists of writing each norm in terms of the inner product, is left to the reader. Putting $A = f - g_n$ and $B = f - g_m$ in the parallelogram law, we find

$$\|2f - (g_n + g_m)\|^2 + \|g_m - g_n\|^2 = 2\left[\|f - g_n\|^2 + \|f - g_m\|^2\right].$$

However $\mathcal{S}$ is a subspace, so the quantity $\frac{1}{2}(g_n + g_m)$ belongs to $\mathcal{S}$, hence

$$\|2f - (g_n + g_m)\| = 2\|f - \frac{1}{2}(g_n + g_m)\| \geq 2d.$$

Therefore

$$\begin{aligned}\|g_m - g_n\|^2 &= 2\left[\|f - g_n\|^2 + \|f - g_m\|^2\right] - \|2f - (g_n + g_m)\|^2 \\ &\leq 2\left[\|f - g_n\|^2 + \|f - g_m\|^2\right] - 4d^2.\end{aligned}$$

By construction, we know that $\|f - g_n\| \to d$ and $\|f - g_m\| \to d$ as $n, m \to \infty$, so the above inequality implies that $\{g_n\}$ is a Cauchy sequence. Since $\mathcal{H}$ is complete and $\mathcal{S}$ closed, the sequence $\{g_n\}$ must have a limit g_0 in $\mathcal{S}$, and then it satisfies $d = \|f - g_0\|$.

We prove that if $g \in \mathcal{S}$, then $g \perp (f - g_0)$. For each ϵ (positive or negative), consider the perturbation of g_0 defined by $g_0 - \epsilon g$. This element belongs to $\mathcal{S}$, hence

$$\|f - (g_0 - \epsilon g)\|^2 \geq \|f - g_0\|^2.$$

Since $\|f - (g_0 - \epsilon g)\|^2 = \|f - g_0\|^2 + \epsilon^2\|g\|^2 + 2\epsilon \operatorname{Re}(f - g_0, g)$, we find that

$$2\epsilon \operatorname{Re}(f - g_0, g) + \epsilon^2\|g\|^2 \geq 0. \tag{5}$$

If $\operatorname{Re}(f - g_0, g) < 0$, then taking ϵ small and positive contradicts (5). If $\operatorname{Re}(f - g_0, g) > 0$, a contradiction also follows by taking ϵ small and negative. Thus $\operatorname{Re}(f - g_0, g) = 0$. By considering the perturbation $g_0 - i\epsilon g$, a similar argument gives $\operatorname{Im}(f - g_0, g) = 0$, and hence $(f - g_0, g) = 0$.

Finally, the uniqueness of g_0 follows from the above observation about orthogonality. Suppose $\tilde{g}_0$ is another point in $\mathcal{S}$ that minimizes the distance to f. By taking $g = g_0 - \tilde{g}_0$ in our last argument we find $(f - g_0) \perp (g_0 - \tilde{g}_0)$, and the Pythagorean theorem gives

$$\|f - \tilde{g}_0\|^2 = \|f - g_0\|^2 + \|g_0 - \tilde{g}_0\|^2.$$

Since by assumption $\|f - \tilde{g}_0\|^2 = \|f - g_0\|^2$, we conclude that $\|g_0 - \tilde{g}_0\| = 0$, as desired.

Using the lemma, we may now introduce a useful concept that is another expression of the notion of orthogonality. If $\mathcal{S}$ is a subspace of a Hilbert space $\mathcal{H}$, we define the **orthogonal complement** of $\mathcal{S}$ by

$$\mathcal{S}^\perp = \{f \in \mathcal{H} : (f, g) = 0 \quad \text{for all } g \in \mathcal{S}\}.$$

Clearly, $S^\perp$ is also a subspace of $\mathcal{H}$, and moreover $\mathcal{S} \cap \mathcal{S}^\perp = \{0\}$. To see this, note that if $f \in \mathcal{S} \cap \mathcal{S}^\perp$, then f must be orthogonal to itself; thus $0 = (f, f) = \|f\|$, and therefore $f = 0$. Moreover, $\mathcal{S}^\perp$ is itself a closed subspace. Indeed, if $f_n \to f$, then $(f_n, g) \to (f, g)$ for every g, by the Cauchy-Schwarz inequality. Hence if $(f_n, g) = 0$ for all $g \in \mathcal{S}$ and all n, then $(f, g) = 0$ for all those g.

Proposition 4.2 *If $\mathcal{S}$ is a closed subspace of a Hilbert space $\mathcal{H}$, then*

$$\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp.$$

The notation in the proposition means that every $f \in \mathcal{H}$ can be written uniquely as $f = g + h$, where $g \in \mathcal{S}$ and $h \in \mathcal{S}^\perp$; we say that $\mathcal{H}$ is the **direct sum** of S and $S^\perp$. This is equivalent to saying that any f in $\mathcal{H}$ is the sum of two elements, one in $\mathcal{S}$, the other in $\mathcal{S}^\perp$, and that $\mathcal{S} \cap \mathcal{S}^\perp$ contains only 0.

The proof of the proposition relies on the previous lemma giving the closest element of f in $\mathcal{S}$. In fact, for any $f \in \mathcal{H}$, we choose g_0 as in the lemma and write

$$f = g_0 + (f - g_0).$$

By construction $g_0 \in \mathcal{S}$, and the lemma implies $f - g_0 \in S^\perp$, and this shows that f is the sum of an element in $\mathcal{S}$ and one in $\mathcal{S}^\perp$. To prove that this decomposition is unique, suppose that

$$f = g + h = \tilde{g} + \tilde{h} \quad \text{ where } g, \tilde{g} \in \mathcal{S} \text{ and } h, \tilde{h} \in \mathcal{S}^\perp.$$

Then, we must have $g - \tilde{g} = \tilde{h} - h$. Since the left-hand side belongs to $\mathcal{S}$ while the right-hand side belongs to $\mathcal{S}^\perp$ the fact that $\mathcal{S} \cap \mathcal{S}^\perp = \{0\}$ implies $g - \tilde{g} = 0$ and $\tilde{h} - h = 0$. Therefore $g = \tilde{g}$ and $h = \tilde{h}$ and the uniqueness is established.

With the decomposition $\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp$ one has the natural projection onto S defined by

$$P_{\mathcal{S}}(f) = g, \quad \text{ where } f = g + h \text{ and } g \in \mathcal{S},\ h \in \mathcal{S}^\perp.$$

The mapping $P_{\mathcal{S}}$ is called the **orthogonal projection** onto $\mathcal{S}$ and satisfies the following simple properties:

(i) $f \mapsto P_{\mathcal{S}}(f)$ is linear,

(ii) $P_{\mathcal{S}}(f) = f$ whenever $f \in \mathcal{S}$,

(iii) $P_{\mathcal{S}}(f) = 0$ whenever $f \in \mathcal{S}^\perp$,

(iv) $\|P_{\mathcal{S}}(f)\| \leq \|f\|$ for all $f \in \mathcal{H}$.

Property (i) means that $P_{\mathcal{S}}(\alpha f_1 + \beta f_2) = \alpha P_{\mathcal{S}}(f_1) + \beta P_{\mathcal{S}}(f_2)$, whenever $f_1, f_2 \in \mathcal{H}$ and α and β are scalars.

It will be useful to observe the following. Suppose $\{e_k\}$ is a (finite or infinite) collection of orthonormal vectors in $\mathcal{H}$. Then the orthogonal projection P in the closure of the subspace spanned by $\{e_k\}$ is given by $P(f) = \sum_k (f, e_k)e_k$. In case the collection is infinite, the sum converges in the norm of $\mathcal{H}$.

We illustrate this with two examples that arise in Fourier analysis.

Example 1. On $L^2([-\pi, \pi])$, recall that if $f(\theta) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ then the partial sums of the Fourier series are

$$S_N(f)(\theta) = \sum_{n=-N}^{N} a_n e^{in\theta}.$$

Therefore, the partial sum operator S_N consists of the projection onto the closed subspace spanned by $\{e_{-N}, \ldots, e_N\}$.

The sum S_N can be realized as a convolution

$$S_N(f)(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(\theta - \varphi) f(\varphi)\, d\varphi,$$

where $D_N(\theta) = \sin((N + 1/2)\theta)/\sin(\theta/2)$ is the **Dirichlet kernel**.

EXAMPLE 2. Once again, consider $L^2([-\pi, \pi])$ and let $\mathcal{S}$ denote the subspace that consists of all $F \in L^2([-\pi, \pi])$ with

$$F(\theta) \sim \sum_{n=0}^{\infty} a_n e^{in\theta}.$$

In other words, $\mathcal{S}$ is the space of square integrable functions whose Fourier coefficients a_n vanish for $n < 0$. From the proof of Fatou's theorem, this implies that $\mathcal{S}$ can be identified with the Hardy space $H^2(\mathbb{D})$, where $\mathbb{D}$ is the unit disc, and so is a closed subspace unitarily isomorphic to $\ell^2(\mathbb{Z}^+)$. Therefore, using this identification, if P denotes the orthogonal projection from $L^2([-\pi, \pi])$ to $\mathcal{S}$, we may also write $P(f)(z)$ for the element corresponding to $H^2(\mathbb{D})$, that is,

$$P(f)(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Given $f \in L^2([-\pi, \pi])$, we define the **Cauchy integral** of f by

$$C(f)(z) = \frac{1}{2\pi i}\int_{\gamma} \frac{f(\zeta)}{\zeta - z}\, d\zeta,$$

where γ denotes the unit circle and z belongs to the unit disc. Then we have the identity

$$P(f)(z) = C(f)(z), \qquad \text{for all } z \in \mathbb{D}.$$

Indeed, since $f \in L^2$ it follows by the Cauchy-Schwarz inequality that $f \in L^1([-\pi, \pi])$, and therefore we may interchange the sum and integral in the following calculation (recall $|z| < 1$):

$$\begin{aligned} P(f)(z) = \sum_{n=0}^{\infty} a_n z^n &= \sum_{n=0}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta} d\theta \right) z^n \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) \sum_{n=0}^{\infty} (e^{-i\theta} z)^n d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(e^{i\theta})}{1 - e^{-i\theta} z} d\theta \\ &= \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{f(e^{i\theta})}{e^{i\theta} - z} i e^{i\theta} d\theta \\ &= C(f)(z). \end{aligned}$$

5 Linear transformations

The focus of analysis in Hilbert spaces is largely the study of their linear transformations. We have already encountered two classes of such transformations, the unitary mappings and the orthogonal projections. There are two other important classes we shall deal with in this chapter in some detail: the "linear functionals" and the "compact operators," and in particular those that are symmetric.

Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are two Hilbert spaces. A mapping $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a **linear transformation** (also called **linear operator** or **operator**) if

$$T(af + bg) = aT(f) + bT(g) \quad \text{for all scalars } a, b \text{ and } f, g \in \mathcal{H}_1.$$

Clearly, linear operators satisfy $T(0) = 0$.

We shall say that a linear operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is **bounded** if there exists $M > 0$ so that

$$\|T(f)\|_{\mathcal{H}_2} \leq M \|f\|_{\mathcal{H}_1}. \tag{6}$$

The **norm** of T is denoted by $\|T\|_{\mathcal{H}_1 \to \mathcal{H}_2}$ or simply $\|T\|$ and defined by

$$\|T\| = \inf M,$$

where the infimum is taken over all M so that (6) holds. A trivial example is given by the **identity operator** I, with $I(f) = f$. It is of course a unitary operator and a projection, with $\|I\| = 1$.

In what follows we shall generally drop the subscripts attached to the norms of elements of a Hilbert space, when this causes no confusion.

Lemma 5.1 $\|T\| = \sup\{|(Tf, g)| : \ \|f\| \leq 1, \ \|g\| \leq 1\}$*, where of course* $f \in \mathcal{H}_1$ *and* $g \in \mathcal{H}_2$.

Proof. If $\|T\| \leq M$, the Cauchy-Schwarz inequality gives

$$|(Tf, g)| \leq M \quad \text{ whenever } \|f\| \leq 1 \text{ and } \|g\| \leq 1;$$

thus $\sup\{|(Tf, g)| : \ \|f\| \leq 1, \ \|g\| \leq 1\} \leq \|T\|$.

Conversely, if $\sup\{|(Tf, g)| : \ \|f\| \leq 1, \ \|g\| \leq 1\} \leq M$, we claim that $\|Tf\| \leq M\|f\|$ for all f. If f or Tf is zero, there is nothing to prove. Otherwise, $f' = f/\|f\|$ and $g' = Tf/\|Tf\|$ have norm 1, so by assumption

$$|(Tf', g')| \leq M.$$

But since $|(Tf', g')| = \|Tf\|/\|f\|$ this gives $\|Tf\| \leq M\|f\|$, and the lemma is proved.

A linear transformation T is **continuous** if $T(f_n) \to T(f)$ whenever $f_n \to f$. Clearly, linearity implies that T is continuous on all of $\mathcal{H}_1$ if and only if it is continuous at the origin. In fact, the conditions of being bounded or continuous are equivalent.

Proposition 5.2 *A linear operator* $T : \mathcal{H}_1 \to \mathcal{H}_2$ *is bounded if and only if it is continuous.*

Proof. If T is bounded, then $\|T(f) - T(f_n)\|_{\mathcal{H}_2} \leq M\|f - f_n\|_{\mathcal{H}_1}$, hence T is continuous. Conversely, suppose that T is continuous but not bounded. Then for each n there exists $f_n \neq 0$ such that $\|T(f_n)\| \geq n\|f_n\|$. The element $g_n = f_n/(n\|f_n\|)$ has norm $1/n$, hence $g_n \to 0$. Since T is continuous at 0, we must have $T(g_n) \to 0$, which contradicts the fact that $\|T(g_n)\| \geq 1$. This proves the proposition.

In the rest of this chapter we shall assume that all linear operators are bounded, hence continuous. It is noteworthy to recall that any linear operator between finite-dimensional Hilbert spaces is necessarily continuous.

5.1 Linear functionals and the Riesz representation theorem

A **linear functional** ℓ is a linear transformation from a Hilbert space $\mathcal{H}$ to the underlying field of scalars, which we may assume to be the complex numbers,

$$\ell : \mathcal{H} \to \mathbb{C}.$$

Of course, we view $\mathbb{C}$ as a Hilbert space equipped with its standard norm, the absolute value.

A natural example of a linear functional is provided by the inner product on $\mathcal{H}$. Indeed, for fixed $g \in \mathcal{H}$, the map

$$\ell(f) = (f, g)$$

is linear, and also bounded by the Cauchy-Schwarz inequality. Indeed,

$$|(f, g)| \leq M\|f\|, \qquad \text{where } M = \|g\|.$$

Moreover, $\ell(g) = M\|g\|$ so we have $\|\ell\| = \|g\|$. The remarkable fact is that this example is exhaustive, in the sense that every *continuous* linear functional on a Hilbert space arises as an inner product. This is the so-called Riesz representation theorem.

Theorem 5.3 *Let ℓ be a continuous linear functional on a Hilbert space $\mathcal{H}$. Then, there exists a unique $g \in \mathcal{H}$ such that*

$$\ell(f) = (f, g) \qquad \textit{for all } f \in \mathcal{H}.$$

Moreover, $\|\ell\| = \|g\|$.

Proof. Consider the subspace of $\mathcal{H}$ defined by

$$\mathcal{S} = \{f \in \mathcal{H} : \ell(f) = 0\}.$$

Since ℓ is continuous the subspace $\mathcal{S}$, which is called the **null-space** of ℓ, is closed. If $\mathcal{S} = \mathcal{H}$, then $\ell = 0$ and we take $g = 0$. Otherwise $\mathcal{S}^\perp$ is non-trivial and we may pick any $h \in \mathcal{S}^\perp$ with $\|h\| = 1$. With this choice of h we determine g by setting $g = \overline{\ell(h)}h$. Thus if we let $u = \ell(f)h - \ell(h)f$, then $u \in \mathcal{S}$, and therefore $(u, h) = 0$. Hence

$$0 = (\ell(f)h - \ell(h)f, h) = \ell(f)(h, h) - (f, \overline{\ell(h)}h).$$

Since $(h, h) = 1$, we find that $\ell(f) = (f, g)$ as desired.

At this stage we record the following remark for later use. Let $\mathcal{H}_0$ be a pre-Hilbert space whose completion is $\mathcal{H}$. Suppose ℓ_0 is a linear functional on $\mathcal{H}_0$ which is bounded, that is, $|\ell_0(f)| \leq M\|f\|$ for all $f \in$

$\mathcal{H}_0$. Then ℓ_0 has an extension ℓ to a bounded linear functional on $\mathcal{H}$, with $|\ell(f)| \leq M\|f\|$ for all $f \in \mathcal{H}$. This extension is also unique. To see this, one merely notes that $\{\ell_0(f_n)\}$ is a Cauchy sequence whenever the vectors $\{f_n\}$ belong to $\mathcal{H}_0$, and $f_n \to f$ in $\mathcal{H}$, as $n \to \infty$. Thus we may define $\ell(f)$ as $\lim_{n\to\infty} \ell_0(f_n)$. The verification of the asserted properties of ℓ is then immediate. (This result is a special case of the extension Lemma 1.3 in the next chapter.)

5.2 Adjoints

The first application of the Riesz representation theorem is to determine the existence of the "adjoint" of a linear transformation.

Proposition 5.4 *Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear transformation. There exists a unique bounded linear transformation T^* on $\mathcal{H}$ so that:*

(i) $(Tf, g) = (f, T^*g)$,

(ii) $\|T\| = \|T^*\|$,

(iii) $(T^*)^* = T$.

The linear operator $T^* : \mathcal{H} \to \mathcal{H}$ satisfying the above conditions is called the **adjoint** of T.

To prove the existence of an operator satisfying (i) above, we observe that for each fixed $g \in \mathcal{H}$, the linear functional $\ell = \ell_g$, defined by

$$\ell(f) = (Tf, g),$$

is bounded. Indeed, since T is bounded one has $\|Tf\| \leq M\|f\|$; hence the Cauchy-Schwarz inequality implies that

$$|\ell(f)| \leq \|Tf\|\,\|g\| \leq B\|f\|,$$

where $B = M\|g\|$. Consequently, the Riesz representation theorem guarantees the existence of a unique $h \in \mathcal{H}$, $h = h_g$, such that

$$\ell(f) = (f, h).$$

Then we define $T^*g = h$, and note that the association $T^* : g \mapsto h$ is linear and satisfies (i).

The fact that $\|T\| = \|T^*\|$ follows at once from (i) and Lemma 5.1:

$$\begin{aligned}\|T\| &= \sup\{|(Tf, g)| : \ \|f\| \leq 1,\ \|g\| \leq 1\}\\ &= \sup\{|(f, T^*g)| : \ \|f\| \leq 1,\ \|g\| \leq 1\}\\ &= \|T^*\|.\end{aligned}$$

To prove (iii), note that $(Tf, g) = (f, T^*g)$ for all f and g if and only if $(T^*f, g) = (f, Tg)$ for all f and g, as one can see by taking complex conjugates and reversing the roles of f and g.

We record here a few additional remarks.

(a) In the special case when $T = T^*$ (we say that T is **symmetric**), then

$$\|T\| = \sup\{|(Tf, f)| : \ \|f\| = 1\}. \tag{7}$$

This should be compared to Lemma 5.1, which holds for any linear operator. To establish (7), let $M = \sup\{|(Tf, f)| : \ \|f\| = 1\}$. By Lemma 5.1 it is clear that $M \leq \|T\|$. Conversely, if f and g belong on $\mathcal{H}$, then one has the following "polarization" identity which is easy to verify

$$\begin{aligned}(Tf, g) =& \frac{1}{4}[(T(f+g), f+g) - (T(f-g), f-g) \\ &+ i\,(T(f+ig), f+ig) - i\,(T(f-ig), f-ig)].\end{aligned}$$

For any $h \in \mathcal{H}$, the quantity (Th, h) is real, because $T = T^*$, hence $(Th, h) = (h, T^*h) = (h, Th) = \overline{(Th, h)}$. Consequently

$$\mathrm{Re}(Tf, g) = \frac{1}{4}\left[(T(f+g), f+g) - (T(f-g), f-g)\right].$$

Now $|(Th, h)| \leq M\|h\|^2$, so $|\mathrm{Re}(Tf, g)| \leq \frac{M}{4}\left[\|f+g\|^2 + \|f-g\|^2\right]$, and an application of the parallelogram law (4) then implies

$$|\mathrm{Re}(Tf, g)| \leq \frac{M}{2}[\|f\|^2 + \|g\|^2].$$

So if $\|f\| \leq 1$ and $\|g\| \leq 1$, then $|\mathrm{Re}(Tf, g)| \leq M$. In general, we may replace g by $e^{i\theta}g$ in the last inequality to find that whenever $\|f\| \leq 1$ and $\|g\| \leq 1$, then $|(Tf, g)| \leq M$, and invoking Lemma 5.1 once again gives the result, $\|T\| \leq M$.

(b) Let us note that if T and S are bounded linear transformations of $\mathcal{H}$ to itself, then so is their product TS, defined by $(TS)(f) = T(S(f))$. Moreover we have automatically $(TS)^* = S^*T^*$; in fact, $(TSf, g) = (Sf, T^*g) = (f, S^*T^*g)$.

(c) One can also exhibit a natural connection between linear transformations on a Hilbert space and their associated bilinear forms. Suppose first that T is a bounded operator in $\mathcal{H}$. Define the corresponding bilinear form B by

$$B(f, g) = (Tf, g). \tag{8}$$

Note that B is linear in f and conjugate linear in g. Also by the Cauchy-Schwarz inequality $|B(f,g)| \leq M\|f\|\,\|g\|$, where $M = \|T\|$. Conversely if B is linear in f, conjugate linear in g and satisfies $|B(f,g)| \leq M\|f\|\,\|g\|$, there is a unique linear transformation so that (8) holds with $M = \|T\|$. This can be proved by the argument of Proposition 5.4; the details are left to the reader.

5.3 Examples

Having presented the elementary facts about Hilbert spaces, we now digress to describe briefly the background of some of the early developments of the theory. A motivating problem of considerable interest was that of the study of the "eigenfunction expansion" of a differential operator L. A particular case, that of a Sturm-Liouville operator, arises on an interval $[a,b]$ of $\mathbb{R}$ with L defined by

$$L = \frac{d^2}{dx^2} - q(x),$$

where q is a given real-valued function. The question is then that of expanding an "arbitrary" function in terms of the eigenfunctions φ, that is those functions that satisfy $L(\varphi) = \mu\varphi$ for some $\mu \in \mathbb{R}$. The classical example of this is that of Fourier series, where $L = d^2/dx^2$ on the interval $[-\pi,\pi]$ with each exponential e^{inx} an eigenfunction of L with eigenvalue $\mu = -n^2$.

When made precise in the "regular" case, the problem for L can be resolved by considering an associated "integral operator" T defined on $L^2([a,b])$ by

$$T(f)(x) = \int_a^b K(x,y)f(y)\,dy,$$

with the property that for suitable f,

$$LT(f) = f.$$

It turns out that a key feature that makes the study of T tractable is a certain compactness it enjoys. We now pass to the definitions and elaboration of some of these ideas, and begin by giving two relevant illustrations of classes of operators on Hilbert spaces.

Infinite diagonal matrix

Suppose $\{\varphi_k\}_{k=1}^{\infty}$ is an orthonormal basis of $\mathcal{H}$. Then, a linear transformation $T : \mathcal{H} \to \mathcal{H}$ is said to be **diagonalized** with respect to the basis

$\{\varphi_k\}$ if

$$T(\varphi_k) = \lambda_k \varphi_k, \qquad \text{where } \lambda_k \in \mathbb{C} \text{ for all } k.$$

In general, a *non-zero* element φ is called an **eigenvector** of T with **eigenvalue** λ if $T\varphi = \lambda\varphi$. So the φ_k above are eigenvectors of T, and the numbers λ_k are the corresponding eigenvalues.

So if

$$f \sim \sum_{k=1}^{\infty} a_k \varphi_k \qquad \text{then} \qquad Tf \sim \sum_{k=1}^{\infty} a_k \lambda_k \varphi_k.$$

The sequence $\{\lambda_k\}$ is called the **multiplier sequence** corresponding to T.

In this case, one can easily verify the following facts:

- $\|T\| = \sup_k |\lambda_k|$.
- T^* corresponds to the sequence $\{\overline{\lambda}_k\}$; hence $T = T^*$ if and only if the λ_k are real.
- T is unitary if and only if $|\lambda_k| = 1$ for all k.
- T is an orthogonal projection if and only if $\lambda_k = 0$ or 1 for all k.

As a particular example, consider $\mathcal{H} = L^2([-\pi, \pi])$, and assume that every $f \in L^2([-\pi, \pi])$ is extended to $\mathbb{R}$ by periodicity, so that $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. Let $\varphi_k(x) = e^{ikx}$ for $k \in \mathbb{Z}$. For a fixed $h \in \mathbb{R}$ the operator U_h defined by

$$U_h(f)(x) = f(x + h)$$

is unitary with $\lambda_k = e^{ikh}$. Hence

$$U_h(f) \sim \sum_{k=-\infty}^{\infty} a_k \lambda_k e^{ikx} \qquad \text{if} \qquad f \sim \sum_{k=-\infty}^{\infty} a_k e^{ikx}.$$

Integral operators, and in particular, Hilbert-Schmidt operators

Let $\mathcal{H} = L^2(\mathbb{R}^d)$. If we can define an operator $T : \mathcal{H} \to \mathcal{H}$ by the formula

$$T(f)(x) = \int_{\mathbb{R}^d} K(x, y) f(y)\, dy \qquad \text{whenever } f \in L^2(\mathbb{R}^d),$$

we say that the operator T is an **integral operator** and K is its associated **kernel**.

In fact, it was the problem of invertibility related to such operators, and more precisely the question of solvability of the equation $f - Tf = g$ for given g, that initiated the study of Hilbert spaces. These equations were then called "integral equations."

In general a bounded linear transformation cannot be expressed as an (absolutely convergent) integral operator. However, there is an interesting class for which this is possible and which has a number of other worthwhile properties: **Hilbert-Schmidt operators**, those with a kernel K that belongs to $L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

Proposition 5.5 *Let T be a Hilbert-Schmidt operator on $L^2(\mathbb{R}^d)$ with kernel K.*

(i) *If $f \in L^2(\mathbb{R}^d)$, then for almost every x the function $y \mapsto K(x, y)f(y)$ is integrable.*

(ii) *The operator T is bounded from $L^2(\mathbb{R}^d)$ to itself, and*

$$\|T\| \leq \|K\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)},$$

where $\|K\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)}$ is the L^2-norm of K on $\mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{2d}$.

(iii) *The adjoint T^* has kernel $\overline{K(y, x)}$.*

Proof. By Fubini's theorem we know that for almost every x, the function $y \mapsto |K(x, y)|^2$ is integrable. Then, part (i) follows directly from an application of the Cauchy-Schwarz inequality.

For (ii), we make use again of the Cauchy-Schwarz inequality as follows

$$\begin{aligned}\left|\int K(x,y)f(y)\,dy\right| &\leq \int |K(x,y)||f(y)|\,dy \\ &\leq \left(\int |K(x,y)|^2\,dy\right)^{1/2} \left(\int |f(y)|^2\,dy\right)^{1/2}.\end{aligned}$$

Therefore, squaring this and integrating in x yields

$$\begin{aligned}\|Tf\|^2_{L^2(\mathbb{R}^d)} &\leq \int \left(\int |K(x,y)|^2 dy \int |f(y)|^2 dy\right) dx \\ &= \|K\|^2_{L^2(\mathbb{R}^d \times \mathbb{R}^d)} \|f\|^2_{L^2(\mathbb{R}^d)}.\end{aligned}$$

Finally, part (iii) follows by writing out (Tf, g) in terms of a double integral, and then interchanging the order of integration, as is permissible by Fubini's theorem.

Hilbert-Schmidt operators can be defined analogously for the Hilbert space $L^2(E)$, where E is a measurable subset of $\mathbb{R}^d$. We leave it to the reader to formulate an prove the analogue of Proposition 5.5 that holds in this case.

Hilbert-Schmidt operators enjoy another important property: they are compact. We will now discuss this feature in more detail.

6 Compact operators

We shall use the notion of sequential compactness in a Hilbert space $\mathcal{H}$: a set $X \subset \mathcal{H}$ is **compact** if for every sequence $\{f_n\}$ in X, there exists a subsequence $\{f_{n_k}\}$ that converges in the norm to an element *in* X.

Let $\mathcal{H}$ denote a Hilbert space, and B the closed unit ball in $\mathcal{H}$,

$$B = \{f \in \mathcal{H} : \|f\| \leq 1\}.$$

A well-known result in elementary real analysis says that in a finite-dimensional Euclidean space, a closed and bounded set is compact. However, this does not carry over to the infinite-dimensional case. The fact is that in this case the unit ball, while closed and bounded, is not compact. To see this, consider the sequence $\{f_n\} = \{e_n\}$, where the e_n are orthonormal. By the Pythagorean theorem, $\|e_n - e_m\|^2 = 2$ if $n \neq m$, so no subsequence of the $\{e_n\}$ can converge.

In the infinite-dimensional case we say that a linear operator $T : \mathcal{H} \to \mathcal{H}$ is **compact** if the closure of

$$T(B) = \{g \in \mathcal{H} : g = T(f) \text{ for some } f \in B\}$$

is a compact set. Equivalently, an operator T is compact if, whenever $\{f_k\}$ is a bounded sequence in $\mathcal{H}$, there exists a subsequence $\{f_{n_k}\}$ so that Tf_{n_k} converges. Note that a compact operator is automatically bounded.

Note that by what has been said, a linear transformation is in general not compact (take for instance the identity operator!). However, if T is of **finite rank**, which means that its range is finite-dimensional, then it is automatically compact. It turns out that dealing with compact operators provides us with the closest analogy to the usual theorems of (finite-dimensional) linear algebra. Some relevant analytic properties of compact operators are given by the proposition below.

Proposition 6.1 *Suppose T is a bounded linear operator on $\mathcal{H}$.*

(i) *If S is compact on $\mathcal{H}$, then ST and TS are also compact.*

(ii) *If $\{T_n\}$ is a family of compact linear operators with $\|T_n - T\| \to 0$ as n tends to infinity, then T is compact.*

(iii) *Conversely, if T is compact, there is a sequence $\{T_n\}$ of operators of finite rank such that $\|T_n - T\| \to 0$.*

(iv) *T is compact if and only if T^* is compact.*

Proof. Part (i) is immediate. For part (ii) we use a diagonalization argument. Suppose $\{f_k\}$ is a bounded sequence in $\mathcal{H}$. Since T_1 is compact, we may extract a subsequence $\{f_{1,k}\}_{k=1}^{\infty}$ of $\{f_k\}$ such that $T_1(f_{1,k})$ converges. From $\{f_{1,k}\}$ we may find a subsequence $\{f_{2,k}\}_{k=1}^{\infty}$ such that $T_2(f_{2,k})$ converges, and so on. If we let $g_k = f_{k,k}$, then we claim $\{T(g_k)\}$ is a Cauchy sequence. We have

$$\|T(g_k) - T(g_\ell)\| \le \|T(g_k) - T_m(g_k)\| + \|T_m(g_k) - T_m(g_\ell)\| + \\ + \|T_m(g_\ell) - T(g_\ell)\|.$$

Since $\|T - T_m\| \to 0$ and $\{g_k\}$ is bounded, we can make the first and last term each $< \epsilon/3$ for some large m independent of k and ℓ. With this fixed m, we note that by construction $\|T_m(g_k) - T_m(g_\ell)\| < \epsilon/3$ for all large k and ℓ. This proves our claim; hence $\{T(g_k)\}$ converges in $\mathcal{H}$.

To prove (iii) let $\{e_k\}_{k=1}^{\infty}$ be a basis of $\mathcal{H}$ and let Q_n be the orthogonal projection on the subspace spanned by the e_k with $k > n$. Then clearly $Q_n(g) \sim \sum_{k>n} a_k e_k$ whenever $g \sim \sum_{k=1}^{\infty} a_k e_k$, and $\|Q_n g\|^2$ is a decreasing sequence that tends to 0 as $n \to \infty$ for any $g \in \mathcal{H}$. We claim that $\|Q_n T\| \to 0$ as $n \to \infty$. If not, there is a $c > 0$ so that $\|Q_n T\| \ge c$, and hence for each n we can find f_n, with $\|f_n\| = 1$ so that $\|Q_n T f_n\| \ge c$. Now by compactness of T, choosing an appropriate subsequence $\{f_{n_k}\}$, we have $Tf_{n_k} \to g$ for some g. But $Q_{n_k}(g) = Q_{n_k} T f_{n_k} - Q_{n_k}(T f_{n_k} - g)$, and hence we conclude that $\|Q_{n_k}(g)\| \ge c/2$, for large k. This contradiction shows that $\|Q_n T\| \to 0$. So if P_n is the complementary projection on the finite-dimensional space spanned by $e_1, \ldots, e_n$, $I = P_n + Q_n$, then $\|Q_n T\| \to 0$ means that $\|P_n T - T\| \to 0$. Since each $P_n T$ is of finite rank, assertion (iii) is established.

Finally, if T is compact the fact that $\|P_n T - T\| \to 0$ implies $\|T^* P_n - T^*\| \to 0$, and clearly $T^* P_n$ is again of finite rank. Thus we need only appeal to the second conclusion to prove the last.

We now state two further observations about compact operators.

- If T can be diagonalized with respect to some basis $\{\varphi_k\}$ of eigenvectors and corresponding eigenvalues $\{\lambda_k\}$, then T is compact if and only if $|\lambda_k| \to 0$. See Exercise 25.
- Every Hilbert-Schmidt operator is compact.

To prove the second point, recall that a Hilbert-Schmidt operator is given on $L^2(\mathbb{R}^d)$ by

$$T(f)(x) = \int_{\mathbb{R}^d} K(x,y)f(y)\,dy, \qquad \text{where } K \in L^2(\mathbb{R}^d \times \mathbb{R}^d).$$

If $\{\varphi_k\}_{k=1}^{\infty}$ denotes an orthonormal basis for $L^2(\mathbb{R}^d)$, then the collection $\{\varphi_k(x)\varphi_\ell(y)\}_{k,\ell\geq 1}$ is an orthonormal basis for $L^2(\mathbb{R}^d \times \mathbb{R}^d)$; the proof of this simple fact is outlined in Exercise 7. As a result

$$K(x,y) \sim \sum_{k,\ell=1}^{\infty} a_{k\ell}\varphi_k(x)\varphi_\ell(y), \qquad \text{with } \textstyle\sum_{k,\ell}|a_{k\ell}|^2 < \infty.$$

We define an operator

$$T_n f(x) = \int_{\mathbb{R}^d} K_n(x,y)f(y)dy, \qquad \text{where } K_n(x,y) = \textstyle\sum_{k,\ell=1}^{n} a_{k\ell}\varphi_k(x)\varphi_\ell(y).$$

Then, each T_n has finite-dimensional range, hence is compact. Moreover,

$$\|K - K_n\|^2_{L^2(\mathbb{R}^d\times\mathbb{R}^d)} = \sum_{k\,\geq\,n \text{ or } \ell\,\geq\,n} |a_{k\ell}|^2 \to 0 \qquad \text{as } n \to \infty.$$

By Proposition 5.5, $\|T - T_n\| \leq \|K - K_n\|_{L^2(\mathbb{R}^d\times\mathbb{R}^d)}$, so we can conclude the proof that T is compact by appealing to Proposition 6.1.

The climax of our efforts regarding compact operators is the infinite-dimensional version of the familiar diagonalization theorem in linear algebra for symmetric matrices. Using a similar terminology, we say that a bounded linear operator T is **symmetric** if $T^* = T$. (These operators are also called "self-adjoint" or "Hermitian.")

Theorem 6.2 (Spectral theorem) *Suppose T is a compact symmetric operator on a Hilbert space $\mathcal{H}$. Then there exists an (orthonormal) basis $\{\varphi_k\}_{k=1}^{\infty}$ of $\mathcal{H}$ that consists of eigenvectors of T. Moreover, if*

$$T\varphi_k = \lambda_k\varphi_k,$$

then $\lambda_k \in \mathbb{R}$ and $\lambda_k \to 0$ as $k \to \infty$.

Conversely, every operator of the above form is compact and symmetric.

The collection $\{\lambda_k\}$ is called the **spectrum** of T.

Lemma 6.3 *Suppose T is a bounded symmetric linear operator on a Hilbert space $\mathcal{H}$.*

(i) *If λ is an eigenvalue of T, then λ is real.*

(ii) *If f_1 and f_2 are eigenvectors corresponding to two distinct eigenvalues, then f_1 and f_2 are orthogonal.*

Proof. To prove (i), we first choose a non-zero eigenvector f such that $T(f) = \lambda f$. Since T is symmetric (that is, $T = T^*$), we find that

$$\lambda(f, f) = (Tf, f) = (f, Tf) = (f, \lambda f) = \overline{\lambda}(f, f),$$

where we have used in the last equality the fact that the inner product is conjugate linear in the second variable. Since $f \neq 0$, we must have $\lambda = \overline{\lambda}$ and hence $\lambda \in \mathbb{R}$.

For (ii), suppose f_1 and f_2 have eigenvalues λ_1 and λ_2, respectively. By the previous argument both λ_1 and λ_2 are real, and we note that

$$\begin{aligned} \lambda_1(f_1, f_2) &= (\lambda_1 f_1, f_2) \\ &= (Tf_1, f_2) \\ &= (f_1, Tf_2) \\ &= (f_1, \lambda_2 f_2) \\ &= \lambda_2(f_1, f_2). \end{aligned}$$

Since by assumption $\lambda_1 \neq \lambda_2$ we must have $(f_1, f_2) = 0$ as desired.

For the next lemma note that every non-zero element of the null-space of $T - \lambda I$ is an eigenvector with eigenvalue λ.

Lemma 6.4 *Suppose T is compact, and $\lambda \neq 0$. Then the dimension of the null space of $T - \lambda I$ is finite. Moreover, the eigenvalues of T form at most a denumerable set $\lambda_1, \ldots, \lambda_k, \ldots$, with $\lambda_k \to 0$ as $k \to \infty$. More specifically, for each $\mu > 0$, the linear space spanned by the eigenvectors corresponding to the eigenvalues λ_k with $|\lambda_k| > \mu$ is finite-dimensional.*

Proof. Let V_λ denote the null-space of $T - \lambda I$, that is, the eigenspace of T corresponding to λ. If V_λ is not finite-dimensional, there exists a countable sequence of orthonormal vectors $\{\varphi_k\}$ in V_λ. Since T is compact, there exists a subsequence $\{\varphi_{n_k}\}$ such that $T(\varphi_{n_k})$ converges.

But since $T(\varphi_{n_k}) = \lambda\varphi_{n_k}$ and $\lambda \neq 0$, we conclude that φ_{n_k} converges, which is a contradiction since $\|\varphi_{n_k} - \varphi_{n_{k'}}\|^2 = 2$ if $k \neq k'$.

The rest of the lemma follows if we can show that for each $\mu > 0$, there are only finitely many eigenvalues whose absolute values are greater than μ. We argue again by contradiction. Suppose there are infinitely many distinct eigenvalues whose absolute values are greater than μ, and let $\{\varphi_k\}$ be a corresponding sequence of eigenvectors. Since the eigenvalues are distinct, we know from the previous lemma that $\{\varphi_k\}$ is orthogonal, and after normalization, we may assume that this set of eigenvectors is orthonormal. One again, since T is compact, we may find a subsequence so that $T(\varphi_{n_k})$ converges, and since

$$T(\varphi_{n_k}) = \lambda_{n_k}\varphi_{n_k}$$

the fact that $|\lambda_{n_k}| > \mu$ leads to a contradiction, since $\{\varphi_k\}$ is an orthonormal set and thus $\|\lambda_{n_k}\varphi_{n_k} - \lambda_{n_j}\varphi_{n_j}\|^2 = \lambda_{n_k}^2 + \lambda_{n_j}^2 \geq 2\mu^2$.

Lemma 6.5 *Suppose $T \neq 0$ is compact and symmetric. Then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. By the observation (7) made earlier, either

$$\|T\| = \sup\{(Tf, f) : \ \|f\| = 1\} \quad \text{or} \quad -\|T\| = \inf\{(Tf, f) : \ \|f\| = 1\}.$$

We assume the first case, that is,

$$\lambda = \|T\| = \sup\{(Tf, f) : \ \|f\| = 1\},$$

and prove that λ is an eigenvalue of T. (The proof of the other case is similar.)

We pick a sequence $\{f_n\} \subset \mathcal{H}$ such that $\|f_n\| = 1$ and $(Tf_n, f_n) \to \lambda$. Since T is compact, we may assume also (by passing to a subsequence of $\{f_n\}$ if necessary) that $\{Tf_n\}$ converges to a limit $g \in \mathcal{H}$. We claim that g is an eigenvector of T with eigenvalue λ. To see this, we first observe that $Tf_n - \lambda f_n \to 0$ because

$$\begin{aligned}
\|Tf_n - \lambda f_n\|^2 &= \|Tf_n\|^2 - 2\lambda(Tf_n, f_n) + \lambda^2\|f_n\|^2 \\
&\leq \|T\|^2\|f_n\|^2 - 2\lambda(Tf_n, f_n) + \lambda^2\|f_n\|^2 \\
&\leq 2\lambda^2 - 2\lambda(Tf_n, f_n) \to 0.
\end{aligned}$$

Since $Tf_n \to g$, we must have $\lambda f_n \to g$, and since T is continuous, this implies that $\lambda Tf_n \to Tg$. This proves that $\lambda g = Tg$. Finally, we must have $g \neq 0$, for otherwise $\|T_n f_n\| \to 0$, hence $(Tf_n, f_n) \to 0$, and $\lambda = \|T\| = 0$, which is a contradiction.

We are now equipped with the necessary tools to prove the spectral theorem. Let $\mathcal{S}$ denote the closure of the linear space spanned by all eigenvectors of T. By Lemma 6.5, the space $\mathcal{S}$ is non-empty. The goal is to prove that $\mathcal{S} = \mathcal{H}$. If not, then since

$$\mathcal{S} \oplus \mathcal{S}^{\perp} = \mathcal{H}, \tag{9}$$

$\mathcal{S}^{\perp}$ would be non-empty. We will have reached a contradiction once we show that $\mathcal{S}^{\perp}$ contains an eigenvector of T. First, we note that T respects the decomposition (9). In other words, if $f \in \mathcal{S}$ then $Tf \in \mathcal{S}$, which follows from the definitions. Also, if $g \in \mathcal{S}^{\perp}$ then $Tg \in \mathcal{S}^{\perp}$. This is because T is symmetric and maps $\mathcal{S}$ to itself, and hence

$$(Tg, f) = (g, Tf) = 0 \quad \text{whenever } g \in \mathcal{S}^{\perp} \text{ and } f \in \mathcal{S}.$$

Now consider the operator T_1, which by definition is the restriction of T to the subspace $S^{\perp}$. The closed subspace $S^{\perp}$ inherits its Hilbert space structure from $\mathcal{H}$. We see immediately that T_1 is also a compact and symmetric operator on this Hilbert space. Moreover, if $S^{\perp}$ is non-empty, the lemma implies that T_1 has a non-zero eigenvector in $S^{\perp}$. This eigenvector is clearly also an eigenvector of T, and therefore a contradiction is obtained. This concludes the proof of the spectral theorem.

Some comments about Theorem 6.2 are in order. If in its statement we drop either of the two assumptions (the compactness or symmetry of T), then T may have no eigenvectors. (See Exercises 32 and 33.) However, when T is a general bounded linear transformation which is symmetric, there is an appropriate extension of the spectral theorem that holds for it. Its formulation and proof require further ideas that are deferred to Chapter 6.

7 Exercises

1. Show that properties (i) and (ii) in the definition of a Hilbert space (Section 2) imply property (iii): the Cauchy-Schwarz inequality $|(f,g)| \leq \|f\| \cdot \|g\|$ and the triangle inequality $\|f+g\| \leq \|f\| + \|g\|$.

[Hint: For the first inequality, consider $(f + \lambda g, f + \lambda g)$ as a positive quadratic function of λ. For the second, write $\|f+g\|^2$ as $(f+g, f+g)$.]

2. In the case of equality in the Cauchy-Schwarz inequality we have the following. If $|(f,g)| = \|f\|\,\|g\|$ and $g \neq 0$, then $f = cg$ for some scalar c.

[Hint: Assume $\|f\| = \|g\| = 1$ and $(f, g) = 1$. Then $f - g$ and g are orthogonal, while $f = f - g + g$. Thus $\|f\|^2 = \|f - g\|^2 + \|g\|^2$.]

3. Note that $\|f + g\|^2 = \|f\|^2 + \|g\|^2 + 2\text{Re}(f, g)$ for any pair of elements in a Hilbert space $\mathcal{H}$. As a result, verify the identity $\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2)$.

4. Prove from the definition that $\ell^2(\mathbb{Z})$ is complete and separable.

5. Establish the following relations between $L^2(\mathbb{R}^d)$ and $L^1(\mathbb{R}^d)$:

(a) Neither the inclusion $L^2(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$ nor the inclusion $L^1(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ is valid.

(b) Note, however, that if f is supported on a set E of finite measure and if $f \in L^2(\mathbb{R}^d)$, applying the Cauchy-Schwarz inequality to $f\chi_E$ gives $f \in L^1(\mathbb{R}^d)$, and

$$\|f\|_{L^1(\mathbb{R}^d)} \leq m(E)^{1/2}\|f\|_{L^2(\mathbb{R}^d)}.$$

(c) If f is bounded ($|f(x)| \leq M$), and $f \in L^1(\mathbb{R}^d)$, then $f \in L^2(\mathbb{R}^d)$ with

$$\|f\|_{L^2(\mathbb{R}^d)} \leq M^{1/2}\|f\|_{L^1(\mathbb{R}^d)}^{1/2}.$$

[Hint: For (a) consider $f(x) = |x|^{-\alpha}$, when $|x| \leq 1$ or when $|x| > 1$.]

6. Prove that the following are dense subspaces of $L^2(\mathbb{R}^d)$.

(a) The simple functions.

(b) The continuous functions of compact support.

7. Suppose $\{\varphi_k\}_{k=1}^{\infty}$ is an orthonormal basis for $L^2(\mathbb{R}^d)$. Prove that the collection $\{\varphi_{k,j}\}_{1\leq k,j<\infty}$ with $\varphi_{k,j}(x, y) = \varphi_k(x)\varphi_j(y)$ is an orthonormal basis of $L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

[Hint: First verify that the $\{\varphi_{k,j}\}$ are orthonormal, by Fubini's theorem. Next, for each j consider $F_j(x) = \int_{\mathbb{R}^d} F(x, y)\overline{\varphi_j(y)}\, dy$. If one assumes that $(F, \varphi_{k,j}) = 0$ for all j, then $\int F_j(x)\overline{\varphi_k(x)}\, dx = 0$.]

8. Let $\eta(t)$ be a fixed strictly positive continuous function on $[a, b]$. Define $\mathcal{H}_\eta = L^2([a, b], \eta)$ to be the space of all measurable functions f on $[a, b]$ such that

$$\int_a^b |f(t)|^2\eta(t)\, dt < \infty.$$

Define the inner product on $\mathcal{H}_\eta$ by

$$(f, g)_\eta = \int_a^b f(t)\overline{g(t)}\eta(t)\, dt.$$

(a) Show that $\mathcal{H}_\eta$ is a Hilbert space, and that the mapping $U : f \mapsto \eta^{1/2} f$ gives a unitary correspondence between $\mathcal{H}_\eta$ and the usual space $L^2([a,b])$.

(b) Generalize this to the case when η is not necessarily continuous.

9. Let $\mathcal{H}_1 = L^2([-\pi,\pi])$ be the Hilbert space of functions $F(e^{i\theta})$ on the unit circle with inner product $(F,G) = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(e^{i\theta})\overline{G(e^{i\theta})}\,d\theta$. Let $\mathcal{H}_2$ be the space $L^2(\mathbb{R})$. Using the mapping

$$x \mapsto \frac{i-x}{i+x}$$

of $\mathbb{R}$ to the unit circle, show that:

(a) The correspondence $U : F \to f$, with

$$f(x) = \frac{1}{\pi^{1/2}(i+x)} F\left(\frac{i-x}{i+x}\right)$$

gives a unitary mapping of $\mathcal{H}_1$ to $\mathcal{H}_2$.

(b) As a result,

$$\left\{\frac{1}{\pi^{1/2}}\left(\frac{i-x}{i+x}\right)^n \frac{1}{i+x}\right\}_{n=-\infty}^{\infty}$$

is an orthonormal basis of $L^2(\mathbb{R})$.

10. Let $\mathcal{S}$ denote a subspace of a Hilbert space $\mathcal{H}$. Prove that $(\mathcal{S}^\perp)^\perp$ is the smallest closed subspace of $\mathcal{H}$ that contains $\mathcal{S}$.

11. Let P be the orthogonal projection associated with a closed subspace $\mathcal{S}$ in a Hilbert space $\mathcal{H}$, that is,

$$P(f) = f \text{ if } f \in \mathcal{S} \quad \text{and} \quad P(f) = 0 \text{ if } f \in \mathcal{S}^\perp.$$

(a) Show that $P^2 = P$ and $P^* = P$.

(b) Conversely, if P is any bounded operator satisfying $P^2 = P$ and $P^* = P$, prove that P is the orthogonal projection for some closed subspace of $\mathcal{H}$.

(c) Using P, prove that if $\mathcal{S}$ is a closed subspace of a separable Hilbert space, then $\mathcal{S}$ is also a separable Hilbert space.

12. Let E be a measurable subset of $\mathbb{R}^d$, and suppose $\mathcal{S}$ is the subspace of $L^2(\mathbb{R}^d)$ of functions that vanish for a.e. $x \notin E$. Show that the orthogonal projection P on $\mathcal{S}$ is given by $P(f) = \chi_E \cdot f$, where χ_E is the characteristic function of E.

13. Suppose P_1 and P_2 are a pair of orthogonal projections on S_1 and S_2, respectively. Then P_1P_2 is an orthogonal projection if and only if P_1 and P_2 commute, that is, $P_1P_2 = P_2P_1$. In this case, P_1P_2 projects onto $S_1 \cap S_2$.

14. Suppose $\mathcal{H}$ and $\mathcal{H}'$ are two completions of a pre-Hilbert space $\mathcal{H}_0$. Show that there is a unitary mapping from $\mathcal{H}$ to $\mathcal{H}'$ that is the identity on $\mathcal{H}_0$.

[Hint: If $f \in \mathcal{H}$, pick a Cauchy sequence $\{f_n\}$ in $\mathcal{H}_0$ that converges to f in $\mathcal{H}$. This sequence will also converge to an element f' in $\mathcal{H}'$. The mapping $f \mapsto f'$ gives the required unitary mapping.]

15. Let T be any linear transformation from $\mathcal{H}_1$ to $\mathcal{H}_2$. If we suppose that $\mathcal{H}_1$ is finite-dimensional, then T is automatically bounded. (If $\mathcal{H}_1$ is not assumed to be finite-dimensional this may fail; see Problem 1 below.)

16. Let $F_0(z) = 1/(1-z)^i$.

(a) Verify that $|F_0(z)| \leq e^{\pi/2}$ in the unit disc, but that $\lim_{r\to 1} F_0(r)$ does not exist.

[Hint: Note that $|F_0(r)| = 1$ and $F_0(r)$ oscillates between ± 1 infinitely often as $r \to 1$.]

(b) Let $\{\alpha_n\}_{n=1}^{\infty}$ be an enumeration of the rationals, and let

$$F(z) = \sum_{j=1}^{\infty} \delta^j F_0(ze^{-i\alpha_j}),$$

where δ is sufficiently small. Show that $\lim_{r\to 1} F(re^{i\theta})$ fails to exist whenever $\theta = \alpha_j$, and hence F fails to have a radial limit for a dense set of points on the unit circle.

17. Fatou's theorem can be generalized by allowing a point to approach the boundary in larger regions, as follows.

For each $0 < s < 1$ and point z on the unit circle, consider the region $\Gamma_s(z)$ defined as the smallest closed convex set that contains z and the closed disc $D_s(0)$. In other words, $\Gamma_s(z)$ consists of all lines joining z with points in $D_s(0)$. Near the point z, the region $\Gamma_s(z)$ looks like a triangle. See Figure 2.

We say that a function F defined in the open unit disc has a **non-tangential limit** at a point z on the circle, if for every $0 < s < 1$, the limit

$$\lim_{\substack{w \to z \\ w \in \Gamma_s(z)}} F(w)$$

exists.

Prove that if F is holomorphic and bounded on the open unit disc, then F has a non-tangential limit for almost every point on the unit circle.

[Hint: Show that the Poisson integral of a function f has non-tangential limits at every point of the Lebesgue set of f.]

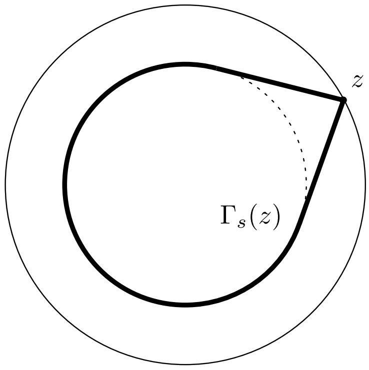

Figure 2. The region $\Gamma_s(z)$

18. Let $\mathcal{H}$ denote a Hilbert space, and $\mathcal{L}(\mathcal{H})$ the vector space of all bounded linear operators on $\mathcal{H}$. Given $T \in \mathcal{L}(\mathcal{H})$, we define the operator norm

$$\|T\| = \inf\{B : \|Tv\| \leq B\|v\|, \quad \text{for all } v \in \mathcal{H}\}.$$

(a) Show that $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$ whenever $T_1, T_2 \in \mathcal{L}(\mathcal{H})$.

(b) Prove that

$$d(T_1, T_2) = \|T_1 - T_2\|$$

defines a metric on $\mathcal{L}(\mathcal{H})$.

(c) Show that $\mathcal{L}(\mathcal{H})$ is complete in the metric d.

19. If T is a bounded linear operator on a Hilbert space, prove that

$$\|TT^*\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2.$$

20. Suppose $\mathcal{H}$ is an infinite-dimensional Hilbert space. We have seen an example of a sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, but for which no subsequence of $\{f_n\}$ converges in $\mathcal{H}$. However, show that for any sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, there exist $f \in \mathcal{H}$ and a subsequence $\{f_{n_k}\}$ such that for all $g \in \mathcal{H}$, one has

$$\lim_{k\to\infty} (f_{n_k}, g) = (f, g).$$

One says that $\{f_{n_k}\}$ **converges weakly** to f.

[Hint: Let g run through a basis for $\mathcal{H}$, and use a diagonalization argument. One can then define f by giving its series expansion with respect to the chosen basis.]

21. There are several senses in which a sequence of bounded operators $\{T_n\}$ can converge to a bounded operator T (in a Hilbert space $\mathcal{H}$). First, there is convergence in the norm, that is, $\|T_n - T\| \to 0$, as $n \to \infty$. Next, there is a weaker convergence, which happens to be called **strong convergence**, that requires that $T_n f \to Tf$, as $n \to \infty$, for every vector $f \in \mathcal{H}$. Finally, there is **weak convergence** (see also Exercise 20) that requires $(T_n f, g) \to (Tf, g)$ for every pair of vectors $f, g \in \mathcal{H}$.

(a) Show by examples that weak convergence does not imply strong convergence, nor does strong convergence imply convergence in the norm.

(b) Show that for any bounded operator T there is a sequence $\{T_n\}$ of bounded operators of finite rank so that $T_n \to T$ strongly as $n \to \infty$.

22. An operator T is an **isometry** if $\|Tf\| = \|f\|$ for all $f \in \mathcal{H}$.

(a) Show that if T is an isometry, then $(Tf, Tg) = (f, g)$ for every $f, g \in \mathcal{H}$. Prove as a result that $T^*T = I$.

(b) If T is an isometry and T is surjective, then T is unitary and $TT^* = I$.

(c) Give an example of an isometry that is not unitary.

(d) Show that if T^*T is unitary then T is an isometry.

[Hint: Use the fact that $(Tf, Tf) = (f, f)$ for f replaced by $f \pm g$ and $f \pm ig$.]

23. Suppose $\{T_k\}$ is a collection of bounded operators on a Hilbert space $\mathcal{H}$, with $\|T_k\| \leq 1$ for all k. Suppose also that

$$T_k T_j^* = T_k^* T_j = 0 \qquad \text{for all } k \neq j.$$

Let $S_N = \sum_{k=-N}^{N} T_k$.

Show that $S_N(f)$ converges as $N \to \infty$, for every $f \in \mathcal{H}$. If $T(f)$ denotes the limit, prove that $\|T\| \leq 1$.

A generalization is given in Problem 8* below.

[Hint: Consider first the case when only finitely many of the T_k are non-zero, and note that the ranges of the T_k are mutually orthogonal.]

24. Let $\{e_k\}_{k=1}^{\infty}$ denote an orthonormal set in a Hilbert space $\mathcal{H}$. If $\{c_k\}_{k=1}^{\infty}$ is a sequence of positive real numbers such that $\sum c_k^2 < \infty$, then the set

$$A = \{\sum_{k=1}^{\infty} a_k e_k : \ |a_k| \leq c_k\}$$

is compact in $\mathcal{H}$.

25. Suppose T is a bounded operator that is diagonal with respect to a basis $\{\varphi_k\}$, with $T\varphi_k = \lambda_k \varphi_k$. Then T is compact if and only if $\lambda_k \to 0$.

[Hint: If $\lambda_k \to 0$, then note that $\|P_n T - T\| \to 0$, where P_n is the orthogonal projection on the subspace spanned by $\varphi_1, \varphi_2, \ldots, \varphi_n$.]

26. Suppose w is a measurable function on $\mathbb{R}^d$ with $0 < w(x) < \infty$ for a.e. x, and K is a measurable function on $\mathbb{R}^{2d}$ that satisfies:

(i) $\displaystyle\int_{\mathbb{R}^d} |K(x,y)| w(y)\, dy \leq A w(x)$ for almost every $x \in \mathbb{R}^d$, and

(ii) $\displaystyle\int_{\mathbb{R}^d} |K(x,y)| w(x)\, dx \leq A w(y)$ for almost every $y \in \mathbb{R}^d$.

Prove that the integral operator defined by

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\, dy, \qquad x \in \mathbb{R}^d$$

is bounded on $L^2(\mathbb{R}^d)$ with $\|T\| \leq A$.

Note as a special case that if $\int |K(x,y)|\, dy \leq A$ for all x, and $\int |K(x,y)|\, dx \leq A$ for all y, then $\|T\| \leq A$.

[Hint: Show that if $f \in L^2(\mathbb{R}^d)$, then

$$\int |K(x,y)|\, |f(y)|\, dy \leq A^{1/2} w(x)^{1/2} \left[\int |K(x,y)|\, |f(y)|^2 w(y)^{-1}\, dy \right]^{1/2}.]$$

27. Prove that the operator

$$Tf(x) = \frac{1}{\pi} \int_0^\infty \frac{f(y)}{x+y}\, dy$$

is bounded on $L^2(0,\infty)$ with norm $\|T\| \leq 1$.

[Hint: Use Exercise 26 with an appropriate w.]

28. Suppose $\mathcal{H} = L^2(B)$, where B is the unit ball in $\mathbb{R}^d$. Let $K(x,y)$ be a measurable function on $B \times B$ that satisfies $|K(x,y)| \leq A|x-y|^{-d+\alpha}$ for some $\alpha > 0$, whenever $x, y \in B$. Define

$$Tf(x) = \int_B K(x,y) f(y) dy.$$

(a) Prove that T is a bounded operator on $\mathcal{H}$.

(b) Prove that T is compact.

(c) Note that T is a Hilbert-Schmidt operator if and only if $\alpha > d/2$.

[Hint: For (b), consider the operators T_n associated with the truncated kernels $K_n(x,y) = K(x,y)$ if $|x-y| \geq 1/n$ and 0 otherwise. Show that each T_n is compact, and that $\|T_n - T\| \to 0$ as $n \to \infty$.]

29. Let T be a compact operator on a Hilbert space $\mathcal{H}$, and assume $\lambda \neq 0$.

(a) Show that the range of $\lambda I - T$ defined by

$$\{g \in \mathcal{H} : g = (\lambda I - T)f, \text{ for some } f \in \mathcal{H}\}$$

is closed. [Hint: Suppose $g_j \to g$, where $g_j = (\lambda I - T)f_j$. Let V_λ denote the eigenspace of T corresponding to λ, that is, the kernel of $\lambda I - T$. Why can one assume that $f_j \in V_\lambda^\perp$? Under this assumption prove that $\{f_j\}$ is a bounded sequence.]

(b) Show by example that this may fail when $\lambda = 0$.

(c) Show that the range of $\lambda I - T$ is all of $\mathcal{H}$ if and only if the null-space of $\overline{\lambda} I - T^*$ is trivial.

30. Let $\mathcal{H} = L^2([-\pi, \pi])$ with $[-\pi, \pi]$ identified as the unit circle. Fix a bounded sequence $\{\lambda_n\}_{n=-\infty}^{\infty}$ of complex numbers, and define an operator Tf by

$$Tf(x) \sim \sum_{n=-\infty}^{\infty} \lambda_n a_n e^{inx} \qquad \text{whenever} \qquad f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$

Such an operator is called a **Fourier multiplier operator**, and the sequence $\{\lambda_n\}$ is called the **multiplier sequence**.

(a) Show that T is a bounded operator on $\mathcal{H}$ and $\|T\| = \sup_n |\lambda_n|$.

(b) Verify that T commutes with translations, that is, if we define $\tau_h(x) = f(x - h)$ then

$$T \circ \tau_h = \tau_h \circ T \qquad \text{for every } h \in \mathbb{R}.$$

(c) Conversely, prove that if T is any bounded operator on $\mathcal{H}$ that commutes with translations, then T is a Fourier multiplier operator. [Hint: Consider $T(e^{inx})$.]

31. Consider a version of the sawtooth function defined on $[-\pi, \pi)$ by[5]

$$K(x) = i(sgn(x)\pi - x),$$

and extended to $\mathbb{R}$ with period 2π. Suppose $f \in L^1([-\pi, \pi])$ is extended to $\mathbb{R}$ with period 2π, and define

$$\begin{aligned} Tf(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} K(x - y) f(y)\, dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} K(y) f(x - y)\, dy. \end{aligned}$$

[5] The symbol $sgn(x)$ denotes the sign function: it equals 1 or -1 if x is positive or negative respectively, and 0 if $x = 0$.

(a) Show that $F(x) = Tf(x)$ is absolutely continuous, and if $\int_{-\pi}^{\pi} f(y)dy = 0$, then $F'(x) = if(x)$ a.e. x.

(b) Show that the mapping $f \mapsto Tf$ is compact and symmetric on $L^2([-\pi, \pi])$.

(c) Prove that $\varphi(x) \in L^2([-\pi, \pi])$ is an eigenfunction for T if and only if $\varphi(x)$ is (up to a constant multiple) equal to e^{inx} for some integer $n \neq 0$ with eigenvalue $1/n$, or $\varphi(x) = 1$ with eigenvalue 0.

(d) Show as a result that $\{e^{inx}\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2([-\pi, \pi])$.

Note that in Book I, Chapter 2, Exercise 8, it is shown that the Fourier series of K is

$$K(x) \sim \sum_{n\neq 0} \frac{e^{inx}}{n}.$$

32. Consider the operator $T : L^2([0, 1]) \to L^2([0, 1])$ defined by

$$T(f)(t) = tf(t).$$

(a) Prove that T is a bounded linear operator with $T = T^*$, but that T is not compact.

(b) However, show that T has no eigenvectors.

33. Let $\mathcal{H}$ be a Hilbert space with basis $\{\varphi_k\}_{k=1}^{\infty}$. Verify that the operator T defined by

$$T(\varphi_k) = \frac{1}{k}\,\varphi_{k+1}$$

is compact, but has no eigenvectors.

34. Let K be a Hilbert-Schmidt kernel which is real and symmetric. Then, as we saw, the operator T whose kernel is K is compact and symmetric. Let $\{\varphi_k(x)\}$ be the eigenvectors (with eigenvalues λ_k) that diagonalize T. Then:

(a) $\sum_k |\lambda_k|^2 < \infty$.

(b) $K(x, y) \sim \sum \lambda_k \varphi_k(x)\varphi_k(y)$ is the expansion of K in the basis $\{\varphi_k(x)\varphi_k(y)\}$.

(c) Suppose T is a compact operator which is symmetric. Then T is of Hilbert-Schmidt type if and only if $\sum_n |\lambda_n|^2 < \infty$, where $\{\lambda_n\}$ are the eigenvalues of T counted according to their multiplicities.

35. Let $\mathcal{H}$ be a Hilbert space. Prove the following variants of the spectral theorem.

(a) If T_1 and T_2 are two linear symmetric and compact operators on $\mathcal{H}$ that commute (that is, $T_1T_2 = T_2T_1$), show that they can be diagonalized simultaneously. In other words, there exists an orthonormal basis for $\mathcal{H}$ which consists of eigenvectors for both T_1 and T_2.

(b) A linear operator on $\mathcal{H}$ is **normal** if $TT^* = T^*T$. Prove that if T is normal and compact, then T can be diagonalized.

[Hint: Write $T = T_1 + iT_2$ where T_1 and T_2 are symmetric, compact and commute.]

(c) If U is unitary, and $U = \lambda I - T$, where T is compact, then U can be diagonalized.

8 Problems

1. Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. There exists a linear functional ℓ defined on $\mathcal{H}$ that is not bounded (and hence not continuous).

[Hint: Using the axiom of choice (or one of its equivalent forms), construct an **algebraic basis** of $\mathcal{H}$, $\{e_\alpha\}$; it has the property that every element of $\mathcal{H}$ is uniquely a finite linear combination of the $\{e_\alpha\}$. Select a denumerable collection $\{e_n\}_{n=1}^{\infty}$, and define ℓ to satisfy the requirement that $\ell(e_n) = n\|e_n\|$ for all $n \in \mathbb{N}$.]

2.* The following is an example of a non-separable Hilbert space. We consider the collection of exponentials $\{e^{i\lambda x}\}$ on $\mathbb{R}$, where λ ranges over the real numbers. Let $\mathcal{H}_0$ denote the space of finite linear combinations of these exponentials. For $f, g \in \mathcal{H}_0$, we define the inner product as

$$(f,g) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(x)\overline{g(x)}\,dx.$$

(a) Show that this limit exists, and

$$(f,g) = \sum_{k=1}^{N} a_{\lambda_k}\overline{b_{\lambda_k}}$$

if $f(x) = \sum_{k=1}^{N} a_{\lambda_k} e^{i\lambda_k x}$ and $g(x) = \sum_{k=1}^{N} b_{\lambda_k} e^{i\lambda_k x}$.

(b) With this inner product $\mathcal{H}_0$ is a pre-Hilbert space. Notice that $\|f\| \leq \sup_x |f(x)|$, if $f \in \mathcal{H}_0$, where $\|f\|$ denotes the norm $\langle f, f\rangle^{1/2}$. Let $\mathcal{H}$ be the completion of $\mathcal{H}_0$. Then $\mathcal{H}$ is not separable because $e^{i\lambda x}$ and $e^{i\lambda' x}$ are orthonormal if $\lambda \neq \lambda'$.

A continuous function F defined on $\mathbb{R}$ is called **almost periodic** if it is the uniform limit (on $\mathbb{R}$) of elements in $\mathcal{H}_0$. Such functions can be identified with (certain) elements in the completion $\mathcal{H}$: We have $\mathcal{H}_0 \subset AP \subset \mathcal{H}$, where AP denotes the almost periodic functions.

(c) A continuous function F is in AP if for every $\epsilon > 0$ we can find a length $L = L_\epsilon$ such that any interval $I \subset \mathbb{R}$ of length L contains an "almost period" τ satisfying

$$\sup_x |F(x+\tau) - F(x)| < \epsilon.$$

(d) An equivalent characterization is that F is in AP if and only if every sequence $F(x + h_n)$ of translates of F contains a subsequence that converges uniformly.

3. The following is a direct generalization of Fatou's theorem: if $u(re^{i\theta})$ is harmonic in the unit disc and bounded there, then $\lim_{r\to 1} u(re^{i\theta})$ exists for a.e. θ.

[Hint: Let $a_n(r) = \frac{1}{2\pi}\int_0^{2\pi} u(re^{i\theta})e^{-in\theta}\,d\theta$. Then $a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{r^2}a_n(r) = 0$, hence $a_n(r) = A_n r^n + B_n r^{-n}$, $n \neq 0$, and as a result[6] $u(re^{i\theta}) = \sum_{-\infty}^{\infty} a_n r^{|n|} e^{in\theta}$. From this one can proceed as in the proof of Theorem 3.3.]

4.* This problem provides some examples of functions that fail to have radial limits almost everywhere.

(a) At almost every point of the boundary unit circle, the function $\sum_{n=0}^{\infty} z^{2^n}$ fails to have a radial limit.

(b) More generally, suppose $F(z) = \sum_{n=0}^{\infty} a_n z^{2^n}$. Then, if $\sum |a_n|^2 = \infty$ the function F fails to have radial limits at almost every boundary point. However, if $\sum |a_n|^2 < \infty$, then $F \in H^2(\mathbb{D})$, and we know by the proof of Theorem 3.3 that F does have radial limits almost everywhere.

5.* Suppose F is holomorphic in the unit disc, and

$$\sup_{0\le r<1} \frac{1}{2\pi}\int_{-\pi}^{\pi} \log^+ |F(re^{i\theta})|\,d\theta < \infty,$$

where $\log^+ u = \log u$ if $u \geq 1$, and $\log^+ u = 0$ if $u < 1$.

Then $\lim_{r\to 1} F(re^{i\theta})$ exists for almost every θ.

The above condition is satisfied whenever (say)

$$\sup_{0\le r<1} \frac{1}{2\pi}\int_{-\pi}^{\pi} |F(re^{i\theta})|^p\,d\theta < \infty, \qquad \text{for some } p > 0,$$

(since $e^{pu} \geq pu$, $u \geq 0$).

Functions that satisfy the latter condition are said to belong to the Hardy space $H^p(\mathbb{D})$.

6.* If T is compact, and $\lambda \neq 0$, show that

[6] See also Section 5, Chapter 2 in Book I.

(a) $\lambda I - T$ is injective if and only if $\overline{\lambda} I - T^*$ is injective.

(b) $\lambda I - T$ is injective if and only if $\lambda I - T$ is surjective.

This result, known as the **Fredholm alternative**, is often combined with that in Exercise 29.

7. Show that the identity operator on $L^2(\mathbb{R}^d)$ cannot be given as an (absolutely) convergent integral operator. More precisely, if $K(x, y)$ is a measurable function on $\mathbb{R}^d \times \mathbb{R}^d$ with the property that for each $f \in L^2(\mathbb{R}^d)$, the integral $T(f)(x) = \int_{\mathbb{R}^d} K(x, y)f(y)\,dy$ converges for almost every x, then $T(f) \neq f$ for some f.

[Hint: Prove that otherwise for any pair of disjoint balls B_1 and B_2 in $\mathbb{R}^d$, we would have that $K(x, y) = 0$ for a.e. $(x, y) \in B_1 \times B_2$.]

8.* Suppose $\{T_k\}$ is a collection of bounded opeartors on a Hilbert space $\mathcal{H}$. Assume that

$$\|T_k T_j^*\| \leq a_{k-j} \quad \text{and} \quad \|T_k^* T_j\| \leq a_{k-j}^*,$$

for positive constants $\{a_n\}$ with the property that $\sum_{-\infty}^{\infty} a_n = A < \infty$. Then $S_N(f)$ converges as $N \to \infty$, for every $f \in \mathcal{H}$, with $S_N = \sum_{-N}^{N} T_k$. Moreover, $T = \lim_{N\to\infty} S_N$ satisfies $\|T\| \leq A$.

9. A discussion of a class of regular Sturm-Liouville operators follows. Other special examples are given in the problems below.

Suppose $[a, b]$ is a bounded interval, and L is defined on functions f that are twice continuously differentiable in $[a, b]$ (we write, $f \in C^2([a, b])$) by

$$L(f)(x) = \frac{d^2 f}{dx^2} - q(x)f(x).$$

Here the function q is continuous and real-valued on $[a, b]$, and we assume for simplicity that q is non-negative. We say that $\varphi \in C^2([a, b])$ is an **eigenfunction** of L with eigenvalue μ if $L(\varphi) = \mu\varphi$, under the assumption that φ satisfies the boundary conditions $\varphi(a) = \varphi(b) = 0$. Then one can show:

(a) The eigenvalues μ are strictly negative, and the eigenspace corresponding to each eigenvalue is one-dimensional.

(b) Eigenvectors corresponding to distinct eigenvalues are orthogonal in $L^2([a, b])$.

(c) Let $K(x, y)$ be the "Green's kernel" defined as follows. Choose $\varphi_-(x)$ to be a solution of $L(\varphi_-) = 0$, with $\varphi_-(a) = 0$ but $\varphi_-'(a) \neq 0$. Similarly, choose $\varphi_+(x)$ to be a solution of $L(\varphi_+) = 0$ with $\varphi_+(b) = 0$, but $\varphi_+'(b) \neq 0$. Let $w = \varphi_+'(x)\varphi_-(x) - \varphi_-'(x)\varphi_+(x)$, be the "Wronskian" of these solutions, and note that w is a non-zero constant.

Set

$$K(x, y) = \begin{cases} \frac{\varphi_-(x)\varphi_+(y)}{w} & \text{if } a \leq x \leq y \leq b, \\ \frac{\varphi_+(x)\varphi_-(y)}{w} & \text{if } a \leq y \leq x \leq b. \end{cases}$$

Then the operator T defined by

$$T(f)(x) = \int_a^b K(x, y)f(y)\,dy$$

is a Hilbert-Schmidt operator, and hence compact. It is also symmetric. Moreover, whenever f is continuous on $[a, b]$, Tf is of class $C^2([a, b])$ and

$$L(Tf) = f.$$

(d) As a result, each eigenvector of T (with eigenvalue λ) is an eigenvector of L (with eigenvalue $\mu = 1/\lambda$). Hence Theorem 6.2 proves the completeness of the orthonormal set arising from normalizing the eigenvectors of L.

10.* Let L be defined on $C^2([-1, 1])$ by

$$L(f)(x) = (1 - x^2)\frac{d^2 f}{dx^2} - 2x\frac{df}{dx}.$$

If φ_n is the n^{th} Legendre polynomial, given by

$$\varphi_n(x) = \left(\frac{d}{dx}\right)^n (1 - x^2)^n, \qquad n = 0, 1, 2, \ldots,$$

then $L\varphi_n = -n(n + 1)\varphi_n$.

When normalized the φ_n form an orthonormal basis of $L^2([-1, 1])$ (see also Problem 2, Chapter 3 in Book I, where φ_n is denoted by L_n.)

11.* The **Hermite functions** $h_k(x)$ are defined by the generating identity

$$\sum_{k=0}^{\infty} h_k(x)\frac{t^k}{k!} = e^{-(x^2/2 - 2tx + t^2)}.$$

(a) They satisfy the "creation" and "annihilation" identities $\left(x - \frac{d}{dx}\right) h_k(x) = h_{k+1}(x)$ and $\left(x + \frac{d}{dx}\right) h_k(x) = h_{k-1}(x)$ for $k \geq 0$ where $h_{-1}(x) = 0$. Note that $h_0(x) = e^{-x^2/2}$, $h_1(x) = 2xe^{-x^2/2}$, and more generally $h_k(x) = P_k(x)e^{-x^2/2}$, where P_k is a polynomial of degree k.

(b) Using (a) one sees that the h_k are eigenvectors of the operator $L = -d^2/dx^2 + x^2$, with $L(h_k) = \lambda_k h_k$, where $\lambda_k = 2k + 1$. One observes that these functions are mutually orthogonal. Since

$$\int_{\mathbb{R}} [h_k(x)]^2\,dx = \pi^{1/2} 2^k k! = c_k,$$

we can normalize them obtaining a orthonormal sequence $\{H_k\}$, with $H_k = c_k^{-1/2} h_k$. This sequence is complete in $L^2(\mathbb{R}^d)$ since $\int_{\mathbb{R}} fH_k\,dx = 0$ for all k implies $\int_{-\infty}^{\infty} f(x)e^{-\frac{x^2}{2} + 2tx}\,dx = 0$ for all $t \in \mathbb{C}$.

(c) Suppose that $K(x,y) = \sum_{k=0}^{\infty} \frac{H_k(x)H_k(y)}{\lambda_k}$, and also $F(x) = T(f)(x) = \int_{\mathbb{R}} K(x,y)f(y)\,dy$. Then T is a symmetric Hilbert-Schmidt operator, and if $f \sim \sum_{k=0}^{\infty} a_k H_k$, then $F \sim \sum_{k=0}^{\infty} \frac{a_k}{\lambda_k} H_k$.

One can show on the basis of (a) and (b) that whenever $f \in L^2(\mathbb{R})$, not only is $F \in L^2(\mathbb{R})$, but also $x^2F(x) \in L^2(\mathbb{R})$. Moreover, F can be corrected on a set of measure zero, so it is continuously differentiable, F' is absolutely continuous, and $F'' \in L^2(\mathbb{R})$. Finally, the operator T is the inverse of L in the sense that

$$LT(f) = LF = -F'' + x^2F = f \qquad \text{for every } f \in L^2(\mathbb{R}).$$

(See also Problem 7* in Chapter 5 of Book I.)