2. The Hilbert space $L^2(S^1)$

We now focus on the class of functions with which Fourier series are most naturally associated. This is the set of "square-summable functions on the circle", or $L^2(S^1)$. Let's start with S^1: this is a circle that has circumference 1, which we can also think of as the interval $[0, 1]$ with the endpoints identified to a single point. A function on the circle is a function of the angle around the circle, or in the interval-with-endpoints-identified model, a function f on $[0, 1]$ for which $f(0) = f(1)$. Another way to think of these functions is as functions defined on all of $\mathbb{R}$ that are periodic with period 1, so that $f(x + 1) = f(x)$ for all x. (There's nothing special about period 1, but using this will make the Fourier series formulas look just like Fourier transform formulas later).

The L^2 part of $L^2(S^1)$ means that the functions we are going to deal with are *complex-valued*, Lebesgue-measurable functions f for which

$$\|f\| = \|f\|_{L^2} = \left(\int_0^1 |f|^2\, dx\right)^{1/2} < \infty.$$

Just as we did last semester, we'll use an analogy with geometry in finite-dimensional space: we'll picture f as a point in an infinite-dimensional space with coordinates $f_x = f(x)$, where x ranges over all the real numbers in $[0, 1]$. If we think of $\|f\|$ as the distance from the "point" f to the origin (the zero function), then the above formula looks like a kind of Pythagorean theorem (length as square root of sum of squares). In this section we'll consider the geometry of $L^2(S^1)$ with this picture in mind.

Important caveat: We have to be careful that we don't end up with two different functions that are at distance zero from one another. To do this, we will declare two functions to be the same if the set of points at which they differ has measure zero (so the Lebesgue integral won't be able to tell them apart). Therefore, f is identified with the zero function if $f = 0$ a.e. This will allow us to assert that $f = 0$ if and only if $\|f\| = 0$. Thus, we are really thinking of *equivalence classes of functions* as the points of $L^2(S^1)$, but we will hardly ever make a big deal about it.

Our goal in this section is to show that $L^2(S^1)$ *is a (separable) Hilbert space*. So we need to say what these words mean.

First, a *Hilbert space* H is a (complex) vector space, meaning it is a set of "points" (or vectors) $\mathbf{v}, \mathbf{w}, \ldots$ that is closed under addition and under multiplication by complex constants (scalar multiplication), subject to all the usual rules: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$, $c(\mathbf{v} + \mathbf{w}) = c\mathbf{v} + c\mathbf{w}$, etc. More than this, H is an *inner product space*, meaning that there is a (Hermitian) inner product on H, that associates a complex number $\langle \mathbf{v}, \mathbf{w} \rangle$ (the inner product, scalar product, or dot product) to any pair of points $\mathbf{v}$ and $\mathbf{w}$ in

H, subject to the rules:

$$\begin{aligned} \langle \mathbf{v}, \mathbf{w} \rangle &= \overline{\langle \mathbf{w}, \mathbf{v} \rangle} \\ \langle \mathbf{v} + \mathbf{w}, \mathbf{x} \rangle &= \langle \mathbf{v}, \mathbf{x} \rangle + \langle \mathbf{w}, \mathbf{x} \rangle \\ \langle c\mathbf{v}, \mathbf{w} \rangle &= c\langle \mathbf{v}, \mathbf{w} \rangle \\ \langle \mathbf{v}, \mathbf{v} \rangle &\geq 0 \end{aligned}$$

and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v}$ is the zero vector. Then, we can define the distance between two points $\mathbf{v}$ and $\mathbf{w}$ to be

$$\|\mathbf{v} - \mathbf{w}\| = \langle \mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w} \rangle^{1/2}.$$

To be a Hilbert space, H is also required to be *complete*. This means that if the distance $\|\mathbf{v}_m - \mathbf{v}_n\| \to 0$ as m and n tend to ∞, then the sequence of points $\{\mathbf{v}_n\}$ must actually converge to some point $\mathbf{w} \in H$ in the sense that

$$\|\mathbf{v}_n - \mathbf{w}\| \to 0$$

as $n \to \infty$ (the content of this statement is that the point $\mathbf{w}$ to which the sequence converges actually exists, in other words, there are no "holes" in the space). Finally, the adjective "separable" means that there is a countable set of points $\{\alpha_n,\ n = 1, 2, \ldots\}$ in H which is *dense*, meaning that any point $\mathbf{v} \in H$ can be approximated arbitrarily closely by a member of the set, i.e.,

$$\inf_{n \geq 1} \|\alpha_n - \mathbf{v}\| = 0.$$

The simplest Hilbert space is the finite-dimensional complex vector space $\mathbb{C}^n$, with points $\mathbf{v} = [v_1, v_2, \ldots, v_n]$, $\mathbf{w} = [w_1, w_2, \ldots, w_n]$, etc., and with inner product defined by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} v_i \overline{w}_i.$$

There is (up to linear isomorphism) only one Hilbert space of each finite dimension, and as we shall see, there is only one infinite-dimensional separable Hilbert space — we can think of it as $L^2(S^1)$, or in a sense as the infinite-dimensional complex vector space $\mathbb{C}^\infty$.

Before we begin to prove that $L^2(S^1)$ is a separable Hilbert space, we'll demonstrate several geometrical facts that are true for all Hilbert spaces, whether finite or infinite-dimensional. So let H be any Hilbert space. We'll start with the fact that when we're dealing with finitely many points in H, it's just like being in $\mathbb{C}^n$, in the sense that all the familiar rules of geometry and trigonometry apply.

First up is the *Schwarz inequality*, which says

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

We proved this last semester for real inner product spaces. To prove this in the complex case, begin with the fact that for any real number ρ and any angle θ, with $0 \leq \theta < 2\pi$, we have that $\mathbf{v} - \rho e^{i\theta}\mathbf{w}$ belongs to H, and that

$$\begin{aligned}\|\mathbf{v} - \rho e^{i\theta}\mathbf{w}\| &= \langle \mathbf{v} - \rho e^{i\theta}\mathbf{w}, \mathbf{v} - \rho e^{i\theta}\mathbf{w}\rangle \\ &= \langle \mathbf{v}, \mathbf{v}\rangle - \langle \rho e^{i\theta}\mathbf{w}, \mathbf{v}\rangle - \langle \mathbf{v}, \rho e^{i\theta}\mathbf{w}\rangle + \langle \rho e^{i\theta}\mathbf{w}, \rho e^{i\theta}\mathbf{w}\rangle \\ &= \|\mathbf{v}\|^2 - 2\,\mathrm{Re}(\rho e^{-i\theta}\langle \mathbf{v}, \mathbf{w}\rangle) + \rho^2\|\mathbf{w}\|^2.\end{aligned}$$

Now choose θ to be the complex argument of $\langle \mathbf{v}, \mathbf{w}\rangle$. Then

$$\|\mathbf{v}\|^2 - 2\rho|\langle \mathbf{v}, \mathbf{w}\rangle| + \rho^2\|\mathbf{w}\|^2 \geq 0,$$

and the left side is a quadratic polynomial in the real variable ρ. Since it is non-negative, its discriminant must be less than or equal to zero. This is:

$$4(|\langle \mathbf{v}, \mathbf{w}\rangle|^2 - \|\mathbf{v}\|^2\|\mathbf{w}\|^2) \leq 0,$$

which is precisely the Schwarz inequality.

Exercise 1: Check that equality holds in Schwarz's inequality if and only if $\mathbf{v}$ and $\mathbf{w}$ are proportional (i.e., $\mathbf{v}$ is a complex scalar multiple of $\mathbf{w}$).

Exercise 2: Show that if the vector $\mathbf{w}$ is fixed, then the inner product $\langle \mathbf{v}, \mathbf{w}\rangle$ is a continuous function of $\mathbf{v}$. Hint: You have to show that $|\langle \mathbf{v}_1, \mathbf{w}\rangle - \langle \mathbf{v}_2, \mathbf{w}\rangle|$ is small if $\|\mathbf{v}_1 - \mathbf{v}_2\|$ is small. Combine the terms and use Schwarz's inequality.

Exercise 3: Prove the *triangle inequality*:

$$\|\mathbf{v} \pm \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|,$$

and interpret it geometrically. Hint for the proof: Show that $\|\mathbf{v} \pm \mathbf{w}\|^2 = \|\mathbf{v}\|^2 \pm 2\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle) + \|\mathbf{w}\|^2$, and use Schwarz's inequality.

Exercise 4: Define the angle between two non-zero vectors $\mathbf{v}$ and $\mathbf{w}$ by:

$$\cos\theta = \frac{\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle)}{\|\mathbf{v}\|\|\mathbf{w}\|}.$$

Explain why you can do this, and then check the "law of cosines":

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos\theta + \|\mathbf{w}\|^2.$$

If we say that $\mathbf{v}$ and $\mathbf{w}$ are perpendicular (or orthogonal) if $\langle \mathbf{v}, \mathbf{w}\rangle = 0$, then the law of cosines becomes the Pythagorean theorem:

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

Prove the generalized Pythagorean theorem:

$$\left\| \sum_{k=1}^{n} \mathbf{v}_k \right\|^2 = \sum_{k=1}^{n} \|\mathbf{v}_k\|^2$$

provided $\langle \mathbf{v}_k, \mathbf{v}_l \rangle = 0$ whenever $k \neq l$.

Now that we've got some elementary geometry down, we can start the proof that $L^2(S^1)$ is a Hilbert space. We'll do this in three steps, the first two of which are pretty easy.

Step 1: $L^2(S^1)$ is closed under multiplication by complex scalars, and under addition. Closure under scalar multiplication is an easy consequence of the linearity of Lebesgue integration. For closure under addition, start with the fact that

$$|f+g|^2 \leq 2|f|^2 + 2|g|^2$$

for any complex numbers f and g, and then integrate this over S^1 with $f(x)$ and $g(x)$ replacing f and g.

Step 2: The inner product makes sense on $L^2(S^1)$. In other words, if $\int |f|^2$ and $\int |g|^2$ are finite, then so is $\int f\overline{g}$. To see this, start with the fact that

$$2|f\overline{g}| = 2|f||g| \leq |f|^2 + |g|^2$$

for any complex numbers f and g. Now you can integrate this over S^1 to see that $f\overline{g}$ is a summable function, and so the inner product

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\, dx$$

makes sense. You can check all the algebraic rules for inner products hold. So we automatically get the Schwarz inequality for functions;

$$|\langle f, g \rangle|^2 = \left| \int_0^1 f\overline{g} \right|^2 \leq \|f\|^2 \|g\|^2 = \int_0^1 |f|^2 \int_0^1 |g|^2.$$

Exercise 5: Explain when equality holds in this context (compare with Exercise 1).

Exercise 6: Prove *Minkowski's inequality*:

$$\left(\int |f+g|^2 \right)^{1/2} \leq \left(\int |f|^2 \right)^{1/2} + \left(\int |g|^2 \right)^{1/2}.$$

Step 3: $L^2(S^1)$ is complete. Note that this would *not* be the case if we used the Riemann integral — that was the point of defining the Lebesgue integral. But before we delve into the details of the proof of this, we need a quick discussion of some of the different ways that a sequence of functions can converge.

Two different ways a sequence of functions can converge: The first kind of convergence is the one we've been gearing up for here, namely convergence *in the sense of the distance in* $L^2(S^1)$: The sequence $\{f_n\}$ converges to f in the L^2 sense if and only if

$$\lim_{n\to\infty} \|f - f_n\|^2 = \lim_{n\to\infty} \int_0^1 |f - f_n|^2 = 0.$$

The other kind of convergence we will consider (which seems natural for Lebesgue measurable functions) is *pointwise convergence almost everywhere.* This means that the measure of the set of $x \in S^1$ for which $f_n(x)$ fails to approach $f(x)$ as $n \to \infty$ is zero. Unfortunately, neither type of convergence implies the other, as shown by the following two examples:

Example 1: Let $f_n(x) = \sqrt{n}$ if $0 \le x \le 1/n$, and $f_n(x) = 0$ otherwise. Then the sequence $\{f_n\}$ converges pointwise almost everywhere (except at $x = 0$ in fact) to the zero function, but

$$\int_0^1 |f_n|^2 = 1$$

for every $n \ge 1$. Thus $\{f_n\}$ does not converge to $f = 0$ in $L^2(S^1)$.

Example 2: Let $\{f_n\}$ be defined as follows: Each f_n is the indicator function of a subinterval of $[0, 1]$, with the subintervals chosen so that every point in $[0, 1]$ is in infinitely many of the subintervals, and the lengths of the subintervals approach zero as $n \to \infty$. Because the lengths of the subintervals go to zero, the L^2-norm $\|f_n\|$ will also go to zero. Thus the sequence $\{f_n\}$ will converge to the zero function in the L^2-sense. But since every point x is in infinitely many of the subintervals, we'll have $f_n(x) = 1$ for infinitely many n, meaning that $f_n(x)$ cannot converge to 0.

To achieve this, we'll let the first two subintervals have length $1/2$ — they'll be $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. Then the next four subintervals will have length $1/4$, being $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{2}{4}]$, $[\frac{2}{4}, \frac{3}{4}]$ and $[\frac{3}{4}, 1]$, and so forth.

But despite these two examples, there is a relationship between L^2 convergence and pointwise convergence almost everywhere:

Theorem: *If the sequence* $\{f_n\}$ *converges to* f *in* $L^2(S^1)$*, then you can find an increasing subsequence of the integers,* $n_1 < n_2 < \cdots \to \infty$ *so that* $\{f_{n_k}\}$ *converges to* f *pointwise a.e. as* $k \to \infty$*.*

The proof of this theorem is a nice application of the monotone convergence theorem, combined with a clever trick. First, because $\lim_{n\to\infty} \|f_n - f\| = 0$, we can choose

$n_1 < n_2 < \cdots$ in such a way that

$$\|f_{n_k} - f\| < \frac{1}{2^k}$$

for $k = 1, 2, \ldots$. But then the monotone convergence theorem (actually, Exercise 7 from section 1) implies that

$$\int_0^1 \sum_{k=1}^{\infty} |f_{n_k}(x) - f(x)|^2 = \sum_{k=1}^{\infty} \int_0^1 |f_{n_k}(x) - f(x)|^2 \leq \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 < \infty,$$

so that

$$\sum_{k=1}^{\infty} |f_{n_k}(x) - f(x)|^2 < \infty \quad \text{a.e.},$$

and this implies the kth term of the sum approaches zero a.e., which is what we needed to show.

Exercise 7: Show that if $\{f_n\}$ converges to f in $L^2(S^1)$ and if it also converges to f^* pointwise a.e., then $f = f^*$ a.e.

Exercise 8: Show that if $\{f_n\}$ converges to f pointwise a.e., and if $\sup_n |f_n(x)|^2$ is summable (i.e., its integral is finite), then $\{f_n\}$ also converges to f in $L^2(S^1)$. (Hint: Use the dominated convergence theorem.)

Now we can tackle the proof that $L^2(S^1)$ is complete. The issue is to check that if $\|f_n - f_m\|$ tends to zero as m and n both tend to ∞, then there is actually an L^2-function f such that

$$\lim_{n\to\infty} \|f_n - f\| = 0.$$

The proof of this is a somewhat elaborate combination of the $1/2^k$ trick with the various theorems (Chebyshev's inequality, Borel-Cantelli, and Fatou's Lemma) concerning the Lebesgue integral.

To begin, by passing to a subsequence, we may as well assume that

$$\|f_k - f_j\|^2 \leq \frac{1}{2^k} \qquad \text{whenever } j \geq k.$$

We're going to show that this subsequence converges pointwise almost everywhere.

Define a sequence of sets as follows:

$$B_j = \{x \mid |f_{j+1}(x) - f_j(x)| \geq \frac{1}{2^{j/3}}\}.$$

We're going to apply the Borel-Cantelli lemma to this sequence of sets, so that we can conclude that almost no point belongs to infinitely many of the B_j's, in other words, that the set

$$B = \bigcap_{k=1}^{\infty} \bigcup_{j \geq k} B_j$$

has measure zero.

Exercise 9: Explain why if $x \notin B$, then $f_j(x)$ converges to a finite limit as $j \to \infty$. (Hint: $x \notin B$ means that $|f_{j+1} - f_j| < 1/2^{j/3}$ for all sufficiently large j).

To apply the Borel-Cantelli lemma to the B_j's, first apply the Chebyshev inequality (Exercise 12 from the preceding section) to see that

$$\mu(B_j) \leq 2^{2j/3} \|f_{j+1} - f_j\|^2 \leq \frac{1}{2^{j/3}}.$$

But $\sum 1/2^{j/3}$ is a convergent series, so we can use the Borel-Cantelli lemma to conclude that $\mu(B) = 0$, and B is the set of x's that are in infinitely many of the B_j's. By Exercise 10, the set C of x's for which the sequence $\{f_j(x)\}$ fails to converge to a finite limit is contained in B, we must have $\mu(C) = 0$ as well. Therefore, our sequence of functions $\{f_j\}$ converges pointwise almost everywhere. Let f be the function to which it converges (if $x \in C$, we may as well just define $f(x) = 0$).

To finish the proof of completeness, we must show that the sequence $\{f_j\}$ also converges to f in $L^2(S^1)$. Because the sequence of f_j's converges pointwise almost everywhere to f, we also have

$$|f_j(x) - f(x)|^2 = \lim_{n\to\infty} |f_j(x) - f_n(x)|^2,$$

and we can then use Fatou's lemma to assert:

$$\|f_j - f\|^2 = \int_0^1 \lim_{n\to\infty} |f_j(x) - f_n(x)|^2 \leq \lim \inf_{n\to\infty} \int_0^1 |f_j(x) - f_n(x)|^2 = \lim \inf_{n\to\infty} \|f_j - f_n\|^2 \leq \frac{1}{2^j}.$$

From this we can conclude two things: First, $f \in L^2(S^1)$, because

$$\|f\| \leq \|f_n\| + \|f_n - f\| \leq \|f_n\| + \frac{1}{2^{n/2}} < \infty,$$

and also that $f_n \to f$ in $L^2(S^1)$m since $1/2^j \to 0$ as $j \to \infty$. We've thus constructed the advertised limit of the sequence, and so we've proved that $L^2(S^1)$ is complete.

Now we know that $L^2(S^1)$ is a Hilbert space. Next, we will show that $L^2(S^1)$ is a *separable* Hilbert space. To do this, we must specify a countable collection of functions $K \subset L^2(S^1)$ such that, given any $f \in L^2(S^1)$, we can approximate f arbitrarily closely with an element of K.

The family we are going to choose is the family of functions that are piecewise constant on S^1, that take only rational values, and that jump at only finitely many rational points of S^1 (considered as the interval $[0,1]$). Certainly, this family is countable, since each of its elements is specified by a finite collection of rational numbers. We have to show that we can approximate any L^2 function on S^1 by elements of K.

To see that any function can be approximated arbitrarily well by an element of our family, we'll suppose we're given the function $f \in L^2(S^1)$, and choose a large integer M and then consider the function

$$g(x) = \begin{cases} f(x) & \text{if } |f(x)| < M \\ 0 & \text{if } |f(x) > M \end{cases}$$

By Chebyshev's inequality, we can make $\|g - f\|$ as small as we like by taking M sufficiently large. Next, we'll approximate g by choosing a large integer N and then setting

$$h(x) = \frac{k}{N} \text{ if } \frac{k}{N} \leq g(x) < \frac{k+1}{N}$$

so that $h(x)$ is always within $1/N$ of $g(x)$ and so we can make $\|h - g\|$ as small as we like by taking N sufficiently large.

We're almost there: the image of $h(x)$ consists of only a finite number of rational values, and so we can write $h(x)$ as a finite sum of functions of the form $(k/N) \cdot 1_A$, where A is the measurable set

$$A = \{x \in S^1 \mid \frac{k}{N} \leq g(x) < \frac{k+1}{N}\}.$$

To finish, we'll show that we can approximate 1_A as closely as we like in L^2 by a function from our class K, which will be the indicator function of a finite union of open intervals.

To do this, recall the "constructive definition of the measure of A", namely, as the infimum of the measures of countable (or what is the same, finite) collections of open intervals that cover A. Thus, we can cover A by a set of open intervals whose total length is as close to $\mu(A)$ as we please. And so we will have approximated 1_A by the indicator function of this union of open intervals, which is in K. This completes the proof that $L^2(S^1)$ is separable.

Exercise 11: Show that any step function (element of K) may be approximated as closely as you like in L^2 by a continuous function. Together with the separability proof above, this shows that the continuous functions are dense in L^2.

Exercise 12: Next show that any step function may be approximated as closely as you like in L^2 by a C^∞ (infinitely differentiable) function. To do this, you'll need to start with a C^∞ function f that has the property that $f(x) = 0$ for $x < 0$ and

$f(x) = 1$ for $x \geq \varepsilon$. Then you can translate and rescale this f to achieve the proof. A prototype for f (with $\varepsilon = 1$) is given by

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-\frac{1}{x} e^{\frac{1}{x-1}}} & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

Exercise 13: Show that for any *non-negative* measurable function φ defined on S^1, the set $L^2(S^1, \varphi(x)dx)$ of measurable functions such that

$$\|f\|_\varphi^2 = \int_0^1 |f(x)|^2 \varphi(x) dx < \infty$$

is a (separable) Hilbert space.

Exercise 14: Show that $L^2(\mathbb{Z}^+)$, which is the set of all (complex) sequences $c = [c_1, c_2, c_3, \ldots]$ such that

$$\|c\|^2 = \sum_{n=1}^{\infty} |c_n|^2 < \infty$$

is a (separable) Hilbert space.