3. Bases for $L^2(S^1)$ and Fourier series

Our next task is to put "coordinates" on the space $L^2(S^1)$, just as you do for $\mathbb{R}^n$ or $\mathbb{C}^n$, so that we can show that $L^2(S^1)$ is linearly isomorphic to the space $L^2(\mathbb{Z}^+)$ from Exercise 14 of the preceding section. Our proof of this will show that in fact that up to isomorphism, *there is only one infinite-dimensional, separable Hilbert space.* To do this, we need some familiar ideas from linear algebra, together with one analytical extension of a linear algebra notion.

Recall that we have the inner product on $L^2(S^1)$:

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

and that we say that f and g are *orthogonal* (i.e., perpendicular) if $\langle f, g\rangle = 0$. A family or sequence of functions f_n, $n = 1, 2, \ldots$ is an *orthogonal family* if $\langle f_i, f_j\rangle = 0$ whenever $i \neq j$. If the functions in an orthogonal family are also of unit length $\|f_i\| = 1$, then we have an *orthonormal* (or unit-perpendicular) family.

Example 1: The functions

$$e_n(x) = e^{2\pi i x}, \quad n \in \mathbb{Z}$$

form an orthonormal family in $L^2([0,1])$, or in $L^2(S^1)$.

Exercise 1: Check the assertion of Example 1. Also, show that for *any* orthonormal family $\{e_i\}$, we have

$$\|e_i - e_j\| = \sqrt{2}$$

whenever $i \neq j$. Combining this with Example 1 shows that there on the surface of the "unit sphere" (the set of functions $\{f \mid \|f\| = 1\} \subset L^2(S^1)$) there are infinitely many points at mutual distances $\sqrt{2}$ – this means that the unit sphere in $L^2(S^1)$ is *not compact*, because no subsequence of the e_n's converges in $L^2(S^1)$.

Exercise 2: Show that for any orthonormal family $\{e_n \mid n \geq 1\}$ in $L^2(S^1)$, and for any sequence $c \in L^2(\mathbb{Z}^+)$, the sum $\sum_{n=1}^{\infty} c_n e_n$ converges in $L^2(S^1)$ to a function f. In other words $\|f - \sum_{n=1}^{N} c_n e_n\|$ tends to zero as $N \to \infty$. Also, prove the "Pythagorean theorem" in this case:

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

Definition: A family of functions $\{f_n\}$ *spans* $L^2(S^1)$ if *finite* sums $\sum_{n=1}^{N} c_k f_k$ with complex coefficients are dense in $L^2(S^1)$. In other words, the family spans $L^2(S^1)$ if for any $f \in L^2(S^1)$ you can make

$$\left\| f - \sum_{n=1}^{N} c_n f_n \right\|$$

as small as you like by choosing N sufficiently large and choosing $c_1, \ldots, c_N$ appropriately.

Naturally, the choice of the coefficients depends upon f and the coefficients might have to change as N is increased. As closer approximations to f are obtained by increasing the number of summands, earlier coefficients might need to be adjusted. An important exception to this observation occurs in the case of an orthonormal family. This is crucial to many numerical curve-fitting schemes and least-squares approximations, as well as to Fourier analysis. We summarize it in the following theorem.

Theorem 1: *Suppose $\{e_n \,|\, n \geq 1\}$ is any orthonormal family in $L^2(S^1)$. Then for any $f \in L^2(S^1)$, and any $N \geq 1$, and any complex numbers $c_1, \ldots, c_N$, we have*

$$\left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n \right\| \leq \left\| f - \sum_{n=1}^{N} c_n e_n \right\| .$$

where $\hat{f}(n)$ is the Fourier coefficient:

$$\hat{f}(n) = \langle f, e_n \rangle = \int_0^1 f(x) \overline{e}_n(x)\, dx.$$

We have equality (i.e., the lower bound is attained) if and only if

$$c_n = \hat{f}(n) \quad \text{for all } n \leq N.$$

Proof. This is a calculation, once you've added zero in a clever way:

$$\begin{aligned}
\left\| f - \sum_{n=1}^{N} c_n e_n \right\|^2 &= \left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n + \sum_{n=1}^{N} \hat{f}(n) e_n - \sum_{n=1}^{N} c_n e_n \right\|^2 \\
&= \left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n + \sum_{n=1}^{N} (\hat{f}(n) - c_n) e_n \right\|^2 \\
&= \left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n \right\|^2 \\
&\quad + 2\,\mathrm{Re} \left(\langle f - \sum_{n=1}^{N} \hat{f}(n) e_n, \sum_{m=1}^{N} (\hat{f}(m) - c_m) e_m \rangle \right) + \left\| \sum_{n=1}^{N} (\hat{f}(n) - c_n) e_n \right\|^2
\end{aligned}$$

But the large inner product in the middle term of the last line is a sum of constant multiples of terms that reduce to zero (Prove this!), and so we are left with

$$\left\| f - \sum_{n=1}^{N} c_n e_n \right\|^2 = \left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n \right\|^2 + \left\| \sum_{n=1}^{N} (\hat{f}(n) - c_n) e_n \right\|^2 .$$

This finishes the proof.

Definition: An *orthonormal basis* of $L^2(S^1)$ is an orthonormal family that spans the whole space.

Exercise 3: Check that an orthonormal family is a basis if and only if

$$f = \sum_{n=1}^{\infty} \hat{f}(n) e_n$$

for any $f \in L^2(S^1)$, where the convergence of the sum is L^2-convergence. This is what we shall call the *Fourier series of* f (with respect to the basis $\{e_n\}$).

Exercise 4: Show that for any orthonormal family and any $f \in L^2(S^1)$,

$$0 \leq \left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n \right\|^2 = \|f\|^2 - \sum_{n=1}^{N} |\hat{f}(n)|^2.$$

From this, deduce *Bessel's inequality*:

$$\sum_{n=1}^{\infty} |\hat{f}(n)|^2 \leq \|f\|^2,$$

and conclude that $\{e_n\}$ is a basis if and only if the *Plancharel identity* holds:

$$\sum_{n=1}^{\infty} |\hat{f}(n)|^2 = \|f\|^2$$

for all $f \in L^2(S^1)$.

Exercise 5: Prove that an orthonormal family spans $L^2(S^1)$ if and only if the only function f that is perpendicular to the every member of the family is $f \equiv 0$. (Hint: Apply exercise 4 to $f - \sum \hat{f}(n) e_n$ – what can this be perpendicular to?)

Theorem 2: *$L^2(S^1)$ has an orthonormal basis.*

Proof: This is why we had to prove that L^2 is separable. Because it is, we know that there is a countable dense family of functions, which we will list as $f_1, f_2, \ldots$. We will begin by going through the list, and throwing out any f_n that can be expressed as a linear combination (with complex coefficients) of the f_j's that precede it in the list. We'll call the functions that survive this process $g_1, g_2, \ldots$. Note that any *finite* subset of the g_n's is a linearly independent set.

We now construct an orthonormal family by applying the Gram-Schmidt process to the g's:

$$
\begin{aligned}
e_1 &= \frac{g_1}{\|g_1\|} \\
&\vdots \\
e_n &= \frac{g_n - \sum_{k<n}\langle g_n, e_k\rangle e_k}{\|g_n - \sum_{k<n}\langle g_n, e_k\rangle e_k\|} \\
&\vdots
\end{aligned}
$$

The weeding out process ensures us that none of the lengths we divide by is zero, and it is clear that the e_n's so constructed are orthonormal. And it is easy to see that the family $\{e_n\}$ spans L^2, since each f from the original list can be expressed as a finite (complex) linear combination of the e_n's. This completes the proof.

Exercise 6: Use the following idea to show that L^2 cannot have a *finite* basis – If A and B are disjoint measurable subsets of S^1 of finite measure, then the indicator functions 1_A and 1_B are orthogonal.

Exercise 7: Later on, we will prove that the functions x^n, $n = 0, 1, 2, \ldots$ span $L^2([-1, 1])$. Calculate the first few functions of the corresponding orthonormal basis by applying the Gram-Schmidt process to these. Up to multiplicative constants, these are the *Legendre polynomials*, which play an important role in the solution of PDEs in spherical coordinates. For the record, $e_1 = \sqrt{\frac{1}{2}}$, $e_2 = \sqrt{\frac{3}{2}}x$. Calculate at least two more.

Next up is the Riesz-Fischer theorem.

Theorem 3: *$L^2(S^1)$ and $L^2(\mathbb{Z}^+)$ are isomorphic. That is, there is a one-to-one linear distance-preserving map of $L^2(S^1)$ to $L^2(\mathbb{Z}^+)$.*

To be more specific, if $\{e_n\}$, $n = 1, 2, \ldots$ is any orthonormal basis of $L^2(S^1)$ and if $\hat{f}(n) = \langle f, e_n\rangle$, then the map $f \mapsto \hat{f}$ is an isomorphism.

Proof: By Exercise 4 above, the Fourier coefficients $\hat{f}(n) = \langle f, e_n\rangle$ satisfy the Plancharel identity

$$\|f\|^2 = \sum_{n=1}^{\infty} |\hat{f}(n)|^2.$$

This shows that $f \mapsto \hat{f}$ is a length-preserving map from $L^2(S^1)$ into $L^2(\mathbb{Z}^+)$. The fact that the map is one-to-one and linear is obvious. Finally, the map is surjective onto $L^2(\mathbb{Z}^+)$. In particular, by Exercise 2, if $c \in L^2(\mathbb{Z}^+)$, then $\sum_{n=1}^{\infty} c_n e_n$ converges to a function $f \in L^2(S^1)$, and

$$\begin{aligned}\hat{f}(k) = \langle f, e_k \rangle &= \left\langle \lim_{n\to\infty} \sum_{j=1}^{n} c_j e_j, e_k \right\rangle \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} c_j \langle e_j, e_k \rangle \\ &= c_k\end{aligned}$$

for every $k \geq 1$. In other words, $\hat{f} = c$. This completes the proof.

Exercise 8: Show that the map $f \mapsto \hat{f}$ automatically preserves inner products:

$$\langle f_1, f_2 \rangle = \int_{S^1} f_1 \overline{f}_2 = \langle \hat{f}_1, \hat{f}_2 \rangle = \sum_{n=1}^{\infty} \hat{f}_1(n) \overline{\hat{f}_2(n)}.$$

Some authors call this the *Parseval identity.* (Hint: Use $\|f_1 - f_2\| = \|\widehat{f_1 - f_2}\| = \|\hat{f}_1 - \hat{f}_2\|$.)

Exercise 9: Check that the Haar functions: $e_0^0(x) = 1$ for $0 \leq x \leq 1$,

$$e_n^k(x) = \begin{cases} 2^{n/2} & \text{for} \quad \dfrac{k-1}{2^n} \leq x < \dfrac{k-\frac{1}{2}}{2^n} \\ -2^{n/2} & \text{for} \quad \dfrac{k-\frac{1}{2}}{2^n} \leq x < \dfrac{k}{2^n} \\ 0 & \text{otherwise} \end{cases}$$

defined for all $n \geq 1$ and $1 \leq k \leq 2^n$, form an orthonormal basis of $L^2([0,1])$. Draw the graphs of a few of them. (Hint: To prove that the functions actually span L^2, show that if f is orthogonal to all of them, then the integral $\int_0^x f$ vanishes for every x of the form $k/2^n$ with $1 \leq k \leq 2^n$ and all $n \geq 0$. Thus $\int_B f = 0$ for every measurable set $B \subset [0,1]$.)

To complete this section, here are some exercises concerning *closed subspaces* of L^2. If A is a subspace of L^2, it means that A is closed under addition and under scalar multiplication (this is the algebraic notion of closure). The adjective "closed" in this context is a topological notion — it means that A is also closed under limits in the sense of L^2. In other words, if $\{f_n\}$ is a sequence of elements of A that converges to f in the L^2 sense, then f also belongs to A (one could say that A is complete with respect to the L^2 norm, it's just not what one says).

Exercise 10: Suppose A is a closed subspace of L^2. The *annihilator subspace* A^0 of A is the set of functions in L^2 that are orthogonal to every function in A. Show that A^0 is also a closed subspace of L^2. Is A^0 still closed even if A is not?

Exercise 11: Let A be a closed subspace again, and let f be an arbitrary function in L^2 (which may or may not be in A). Show there is a function that we will call Pf in A which is closest to f, in other words

$$\|f - Pf\| \leq \|f - g\|$$

for any $g \in A$. (Hint: choose a sequence of g_n's whose distances from f approach the infimum of distances over all elements of A. Use the closure of A to do the rest.

Exercise 12: Show that $f - Pf$ belongs to A^0, and show that $f = Pf + (f - Pf)$ is in fact that *only* way of splitting f into a sum of a piece from A and a piece from A^0. This situation is summed up by saying that L^2 is the *orthogonal direct sum* of A and A^0:

$$L^2 = A \oplus A^0.$$

(Hint: Choose any $g \in A$. Then $\varphi(\varepsilon) = \|f - Pf + \varepsilon g\|^2$ is a polynomial of degree 2 and has a minimum for $\varepsilon = 0$. Calculate $\varphi'(0)$.)

Exercise 13: The map $f \mapsto Pf$ is called the *projection* onto A. Show that P is a linear map of L^2 into itself, in other words, it repsects addition and scalar multiplication. Also, show that $P^2 = P$, $\langle Pf_1, f_2 \rangle = \langle f_1, Pf_2 \rangle$, $\|Pf\| \leq \|f\|$, and P is the identity map on A and is zero on A^0. Finally, show that any closed subspace A of L^2 must be separable.

Exercise 14: Show that the family of functions $\{f_n\}$ for $n = 1, \ldots$ spans L^2 if and only if $\langle f, f_n \rangle = 0$ for every $n \geq 1$ implies $f \equiv 0$. (Hint: What is the annihilator of the family $\{f_n\}$?)

Exercise 15: Show that any linear map φ from L^2 into the complex numbers which is bounded in the sense that

$$|\varphi(f)| \leq C\|f\|$$

for some constant C which is independent of f, cna be expressed as an inner product:

$$\varphi(f) = \langle f, g \rangle$$

for some $g \in L^2$. This is called the *Riesz representation theorem*. (Hint: Assume $\varphi \neq 0$ and let A be the set $\{f \mid \varphi(f) = 0\}$. Explain why A is closed and show that A^0 has dimension 1. Then find a function $g \in L^2$ so that $\varphi(f) = 0$ if and only if $\langle f, g \rangle = 0$. Then

$$\varphi(f) = \frac{\varphi(g)\langle f, g \rangle}{\|g\|^2} = \alpha \langle f, g \rangle$$

for some constant α.)