Fourier Analysis

Notes by Albert Zhang

May 5, 2020

Contents

1 Fourier Series, Pre-Lebesgue

Definition 1.1 (Fourier Series)**.** Let $f : [a, b] \to \mathbb{R}$ be an integrable function, where $L = b - a$. The n^{th} Fourier coefficient $\hat{f}(n) = a_n$ is defined as

$$a_n := \frac{1}{L} \int_a^b f(x) e^{-2\pi i n x/L} dx, \quad n \in \mathbb{Z}.$$

The *Fourier series* of f is then given by

$$\sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x/L}.$$

Often, we will consider f on the interval $[-\pi, \pi]$, so that our fourier coefficients become

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta, \quad n \in \mathbb{Z},$$

and the Fourier series of f becomes

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

We define the N^{th} partial sum as

$$S_N(f)(x) = \sum_{n=-N}^{N} a_n e^{2\pi i n x/L},$$

and immediately the question of convergence arises. That is, in what sense does $S_N(f) \to f$? Namely, we will explore pointwise, uniform, and mean convergence in the coming sections.

1.1 Uniqueness & Uniform Convergence

First, we deal with the matter of uniqueness: if f and g have the same Fourier coefficients a_n, then is $f = g$ almost everywhere? In particular, this reduces to showing that $f = 0$ almost everywhere given that $a_n = 0$ for all $n \in \mathbb{Z}$. The following result verifies this for continuous functions on the circle.

Theorem 1.2 (Uniqueness of Fourier Series)

Suppose f is a continuous function on the circle with $a_n = 0$ for $n \in \mathbb{Z}$. Then $f = 0$. More generally, if f is piecewise continuous, then $f = 0$ almost everywhere.

Proof. Suppose WLOG that f is real-valued, defined on $[-\pi, \pi]$, and f is continuous at 0. For sake of contradiction, let $f(0) > 0$. The idea is to construct a sequence of trigonometric polynomials $\{p_k\}$ which blow up near zero, so that

$$\int p_k(\theta) f(\theta) d\theta \to \infty,$$

contradicting the fact that these integrals are all zero, since we can decompose each of them into sums of Fourier coefficients, which we have assumed to be zero.

Let $p(\theta) = \cos\theta + \epsilon$, where we will pick the parameter ϵ later. We approach this by picking parameters $0 < \eta < \delta \leq \pi/2$ such that for $|\theta| < \eta$, our p_k blows up; for $\eta < |\theta| < \delta$, the product $p_k(\theta)f(\theta)$ is nonnegative; and for $\delta \leq |\theta| \leq \pi$, the sequence $|p_k(\theta)|$ vanishes.

More precisely, since f is continuous at 0, we can pick $0 < \delta \leq \pi/2$ such that $f(\theta) > 0$ for $|\theta| < \delta$. Then let

$$p(\theta) = \cos\theta + \epsilon,$$

where $\epsilon > 0$ is chosen so small that $|p(\theta)| < 1 - \epsilon/2$ for $\delta \leq |\theta| \leq \pi$. Then, choose $0 < \eta < \delta$ so that $p(\theta) \geq 1 + \epsilon/1$ for $|\theta| < \eta$. Now, let

$$p_k(\theta) = [p(\theta)]^k.$$

Since f is integrable, and thus bounded, we have the bound

$$\left| \int_{\delta \leq |\theta| \leq \pi} p_k(\theta) f(\theta) \right| \leq 2\pi B(1 - \epsilon/2)^k,$$

for some constant B, so that the integral vanishes as $k \to \infty$. On the other hand, by our choice of δ, we know

$$\int_{\eta \leq |\theta| < \delta} p_k(\theta) f(\theta) d\theta \geq 0.$$

Finally, the integral near 0

$$\int_{|\theta| < \eta} p_k(\theta) f(\theta) d\theta \geq 2\eta \left(\frac{f(0)}{2} \right) \left(1 + \frac{\epsilon}{2} \right)^k.$$

It follows that

$$\int_{-\pi}^{\pi} p_k(\theta) f(\theta) d\theta \to \infty,$$

contradiction. Note for complex-valued functions, we can do the usual and split the function into its real and imaginary parts. Hence f must vanish almost everywhere. □

As a corollary, we get our first uniform convergence type result:

Corollary 1.3

Suppose f is a continuous function on the circle and that its Fourier series is absolutely convergent, i.e.

$$\sum_{n=-\infty}^{\infty} |a_n| < \infty.$$

Then the Fourier series converges uniformly to f:

$$\lim_{N \to \infty} S_N(f) = f.$$

Proof. Note that for $\theta \in [-\pi, \pi]$, we have

$$\left| \sum_{n=-\infty}^{-N} a_n e^{in\theta} + \sum_{n=N}^{\infty} a_n e^{in\theta} \right| \leq \sum_{n=-\infty}^{-N} |a_n| + \sum_{n=N}^{\infty} |a_n| \to 0,$$

since the tail sums of the absolute series must vanish. Then, recall that the uniform limit of a sequence of continuous functions is continuous, so the function

$$g(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta} = \lim_{N\to\infty} \sum_{n=-N}^{N} a_n e^{in\theta}$$

is continuous. Now, we can exchange summation with the integral due to uniform convergence to get

$$\begin{aligned} b_k &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} a_n e^{in\theta} \cdot e^{ik\theta} d\theta \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_{-\pi}^{\pi} a_n e^{i(n-k)\theta} d\theta \\ &= a_k. \end{aligned}$$

Thus, the Fourier coefficients of g are just the Fourier coefficients of f. Now we can apply Theorem 1.2 to deduce that $f = g$, and hence we have uniform convergence of the Fourier series of f. ☐

It turns out that the smoothness of f is directly related to the decay of the Fourier coefficients, and therefore also the uniform convergence of the Fourier series.

Theorem 1.4 (Uniform Convergence in C^2)
Suppose $f \in C^2([0, 2\pi])$ is a periodic function on the circle. Then $a_n = O(1/n^2)$ as $|n| \to \infty$, so that the Fourier series of f converges absolutely and therefore uniformly to f.

Proof. Using integration by parts twice,

$$\begin{aligned} 2\pi a_n &= \int_0^{2\pi} f(\theta) e^{in\theta} d\theta \\ &= \left[f(\theta) \cdot \frac{-e^{-in\theta}}{in} \right]_0^{2\pi} + \frac{1}{in} \int_0^{2\pi} f'(\theta) e^{-in\theta} d\theta \\ &= \frac{1}{in} \int_0^{2\pi} f'(\theta) e^{-in\theta} d\theta \\ &= \frac{1}{in} \left[f'(\theta) \cdot \frac{-e^{-in\theta}}{in} \right]_0^{2\pi} + \frac{1}{(in)^2} \int_0^{2\pi} f''(\theta) e^{-in\theta} d\theta \\ &= \frac{-1}{n^2} \int_0^{2\pi} f''(\theta) e^{-in\theta} d\theta. \end{aligned}$$

Now, since the second derivative is continuous on a compact interval, it is bounded, and hence

$$2\pi n^2|a_n| \leq \left|\int_0^{2\pi} f''(\theta)e^{-in\theta}d\theta\right| \leq \int_0^{2\pi} |f''(\theta)|d\theta \leq C.$$

Thus, $\sum |a_n|$ is dominated by $\sum 1/n^2 < \infty$, and so applying Corollary 1.3 we obtain uniform convergence. $\square$

It turns out uniform convergence can also be shown under the stronger condition of $f \in C^1$. However, our proof fails when we try to modify it to prove that stronger statement.

1.2 Convolution

Definition 1.5 (Convolution)**.** Given two 2π-periodic integrable functions f and g on $\mathbb{R}$, we define their *convolution* $f * g$ on $[-\pi, \pi]$ as

$$(f * g)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)g(x-t)dt.$$

Actually, convolution is usually defined without the normalizing factor in front. However, for the purpose of this section, we will follow Shakarchi's way of definition.

One motivation for convolutions comes from the relationship between the partial sums of Fourier series and the Dirichlet kernels. In particular, the N^{th} Dirichlet kernel D_N is given by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx},$$

and it turns out that

$$\begin{aligned}(f * D_N)(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\left(\sum_{n=-N}^{N} e^{in(x-y)}\right)dy \\ &= \sum_{n=-N}^{N}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}dy\right)e^{inx} \\ &= \sum_{n=-N}^{N} a_n e^{inx} \\ &= S_N(f)(x).\end{aligned}$$

Note the connection to EE: the n^{th} Fourier coefficients are essentially the frequency responses to each complex exponential e^{inx} for a system whose impulse response is given by f.

Proposition 1.6

Suppose f, g, and h are 2π-periodic integrable functions. Then we have the following properties:

(i) $f * (g + h) = (f * g) + (f * h)$.
(ii) $(cf) * g = c(f * g) = f * (cg)$ for any $c \in \mathbb{C}$.
(iii) $f * g = g * f$.
(iv) $(f * g) * h = f * (g * h)$.
(v) $f * g$ is continuous.
(vi) The Fourier coefficients of the convolution are the product of the Fourier coefficients of f and g:

$$\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n).$$

Properties (i)-(iv) are straightforward. Note that (v) says, surprisingly, that $f * g$ is more "regular" than its operands. Property (vi) in EE terms says that the frequency response of the convolution (which is just the composition of systems by associativity) is just the product of the frequency responses of the individual systems, which makes sense if we think of frequency responses as a sort of eigenvalue for the system. We will postpone the proof of these properties for later when we are equipped the Lebesgue integral, which makes things easier.

1.3 L^2 Convergence

Our goal is to prove mean square convergence of a Fourier series. Before we begin, we need the following two lemmas:

Lemma 1.7 (Approximating Integrable by Continuous)

Suppose f is integrable on the circle and bounded by some constant C. Then there exists a sequence (f_n) of continuous functions on the circle uniformly bounded by C so that

$$\int_{-\pi}^{\pi} |f(x) - f_n(x)|dx \to 0 \quad \text{as } n \to \infty.$$

Proof. Actually, we can prove something stronger. We can relax f to be defined on any interval, finite or infinite. Then since f is integrable, there exists a sequence of step functions (given by partitions) which approximate f. We also know that step functions can be approximated by continuous functions, defined by replacing each neighborhood of discontinuity with a slanted line. It's easy to show then, using the $\epsilon/2^n$ trick (the step functions have at most countable many discontinuities), that overall the L^1 distance between the step functions and the continuous functions can be made arbitrarily small. □

Before we move on to the next lemma, we need the concept of a "good" kernel.

We can think of these as an approximation to the identity, or the "distribution" which represents the dirac delta (a point mass at 0).

Definition 1.8 (Good Kernels)**.** We say a family of kernels (K_n) is good if it satisfies

(i) For all $n \geq 1$,
$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)dx = 1.$$

(ii) There exists $M > 0$ such that for all $n \geq 1$,
$$\int_{-\pi}^{\pi} |K_n(x)|dx \leq M.$$

(iii) For every $\delta > 0$,
$$\int_{\delta \leq |x| \leq \pi} |K_n(x)|dx \to 0, \qquad \text{as } n \to \infty.$$

Proposition 1.9

Let (K_n) be a family of good kernels, and f an integrable function on the circle. Then
$$\lim_{n\to\infty} (f * K_n)(x) = f(x)$$
whenever f is continuous at x. If f is continuous everywhere, then the limit is uniform.

Proof. Let $\epsilon > 0$, and f be continuous at x. So choose δ so that $|y| < \delta$ implies $|f(x-y) - f(x)| < \epsilon$. Then though the defining properties of a good kernel, we have

$$\begin{aligned}
|(f * K_n)(x) - f(x)| &= \left| \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(y)(f(x-y) - f(x))dy \right| \\
&\leq \frac{1}{2\pi}\int_{|y|<\delta} |K_n(y)||f(x-y) - f(x)|dy \\
&\quad + \frac{1}{2\pi}\int_{\delta \leq |y| \leq \pi} |K_n(y)||f(x-y) - f(x)|dy \\
&\leq \frac{\epsilon M}{2\pi} + \frac{B}{2\pi}\int_{\delta \leq |y| \leq \pi} |K_n(y)|dy.
\end{aligned}$$

By the vanishing property (iii), we see that the second term goes to zero, and so
$$(f * K_n)(x) \to f(x).$$

If in addition f is continuous on $[-\pi, \pi]$, then it is uniformly continuous, so in particular our choice of δ works for all $x \in [-\pi, \pi]$, giving us uniform convergence. □

Lemma 1.10 (Fejér's Theorem, Approximating Continuous by Trigonometric)
Let f be a continuous function on the circle $[-\pi, \pi]$. Then f can be uniformly approximated by trigonometric polynomials.

Proof. For a sequence of partial sums (s_n) of some series $\sum z_k$, we call σ_N the N^{th} Cesaro mean or sum of the series, given by

$$\sigma_N = \frac{s_0 + s_1 + \cdots + s_{N-1}}{N}.$$

If $\lim \sigma_N = \sigma$ exists, then we say that the series $\sum z_k$ is Cesaro summable to σ. In particular, form the N^{th} Cesaro mean of the Fourier series, given by

$$\sigma_N(f)(x) = \frac{S_0(f)(x) + \cdots + S_{N-1}(f)(x)}{N}.$$

Recall that $S_n(f) = f * D_n$, so it follows that

$$\sigma_N(f)(x) = (f * F_N)(x),$$

where we define $F_N(x)$ to be the N^{th} Fejér kernel given by the average of first N Dirichlet kernels.

$$F_N(x) = \frac{D_0(x) + \cdots + D_{N-1}(x)}{N}.$$

Now, through some basic trig-bashing, we get that

$$F_N(x) = \frac{\sin^2(Nx/2)}{N\sin^2(x/2)}.$$

Now, we show that the Fejér kernel is a good kernel. Note that $F_N(x) \geq 0$, and by some simple calculations (most of the e^{inx} terms of the Dirichlet kernel cancel out),

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_N(x)dx = 1.$$

Now, for $\delta > 0$, we have that $\sin^2(x/2) \geq \sin^2(\delta/2) > 0$ for $\delta \leq |x| \leq \pi$. Therefore $F_N(x) \leq 1/(N\sin^2(\delta/2))$, so taking $N \to \infty$, gives

$$\int_{\delta \leq |x| \leq \pi} |F_N(x)|dx \to 0.$$

Therefore, by Proposition 1.9, we have the uniform limit

$$\lim_{N\to\infty} \sigma_N(f)(x) = \lim_{N\to\infty}(f * F_N)(x) = f(x).$$

Since the Cesaro means are themselves trigonometric polynomial, the claim follows. $\square$

Equipped with the these lemmas, we are ready to prove convergence in mean of a Fourier series.

Theorem 1.11 (Mean-Square Convergence)
Suppose f is integrable on the circle. Then as $N \to \infty$,

$$\frac{1}{2\pi}\int_0^{2\pi} |f(\theta) - S_N(f)(\theta)|^2 d\theta \to 0.$$

Proof. The idea is to first realize the Fourier sums as projections onto a growing orthonormal family of basis functions. In particular, it's easy to verify the family $\mathcal{B}_N := \{e^{inx}\}_{|n|\leq N}$ is orthonormal with respect to the complex inner product given by

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f \cdot \overline{g} dx.$$

Now, note that the Fourier coefficients are just

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

It follows that the partial sums

$$S_N(f)(x) = \sum_{n=-N}^{N} \langle f, e^{inx}\rangle e^{inx}$$

are just projections of f onto the N^{th} hyperplane spanned by the basis $\mathcal{B}_N$. Therefore

$$\|f(x) - S_N(f)(x)\|_2 \leq \|f(x) - \sum_{n=-N}^{N} b_n e^{inx}\|_2,$$

for any set of coefficients $b_n \in \mathbb{C}$. Now, by 1.7 and 1.10, we have that the trigonometric polynomials are dense in the space of continuous functions, which are dense in the space of integrable functions. Therefore, we can approximate any integrable function uniformly by a sequence of trigonometric polynomials p_N formed from the growing bases $\mathcal{B}_N$. Setting $\sum_{n=-N}^{N} b_n e^{inx}$ to this sequence gives us

$$\begin{aligned}\|f(x) - S_N(f)(x)\|_2 &\leq \|f(x) - p_N(x)\|_2 \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - p_N(x)|^2 dx \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \epsilon_N^2 dx,\end{aligned}$$

where $\epsilon_N \to 0$ as $N \to \infty$. This shows that

$$\|f(x) - S_N(f)(x)\|_2 \to 0,$$

as desired. □

Recall that each $S_N(f)(x)$ is the projection of f onto the orthonormal basis $\mathcal{B}_N$. Then we have

$$\|f\|_2^2 = \|f - S_N(f)\|_2^2 + \sum_{n=-N}^{N} |a_n|^2.$$

In particular, if we take $N \to \infty$, then we get Parseval's identity:

Corollary 1.12 (Parseval's Identity)

If f is integrable on the circle, then

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \|f\|_2^2.$$

Since $\sum_{n=-\infty}^{\infty} |a_n|^2$ converges, it follows that its terms go to zero, so we get the following result.

Theorem 1.13 (Riemann-Lebesgue Lemma)

If f is integrable on the circle, then $a_n = \hat{f}(n) \to 0$ as $|n| \to \infty$.

From the Riemann-Lebesgue lemma, it turns out we can obtain another condition for pointwise convergence. Its proof we omit for now, as we wait until we have more convenient tools (i.e. Lebesgue integration).

Corollary 1.14 (A Condition for Pointwise Convergence)

Let f be an integrable function on the circle. If it is differentiable or Lipschitz continuous at x_0, then $S_N(f)(x_0) \to f(x_0)$ as $N \to \infty$.

2 Fourier Series

2.1 Tools of Lebesgue

From now on, we will assume familiarity with basic measure theory and the Lebesgue integral. The following tools will be crucial:

Theorem 2.1 (Monotone Convergence)
Suppose $0 \leq f_n \uparrow f$ are measurable functions. Then

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

The condition $0 \leq f_1$ may be replaced by the condition that $\int f_1^- < \infty$.

Theorem 2.2 (Dominated Convergence)
Suppose $|f_n| < g$, where g is absolutely integrable. If $f_n \to f$, thn

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Lemma 2.3 (Fatou)
If $f_n \geq 0$, then

$$\liminf_{n\to\infty} \int f_n d\mu \geq \int \liminf_{n\to\infty} f_n d\mu.$$

Lemma 2.4 (Borel-Cantelli)
Let $B_n \subset \Omega$ for $n \geq 1$. If $\sum_{n=1}^{\infty} \mu(B_n) < \infty$, then

$$\mu(\omega \in B_n \text{ i.o.}) := \mu(\cap_{n\geq 1} \cup_{m\geq n} B_m) = 0.$$

Theorem 2.5 (Fubini)
Let f be a measurable function defined on the product space $(\Omega, \Sigma_1 \times \Sigma_2, \mu_1 \times \mu_2) := (\Omega_1, \Sigma_1, \mu_1) \times (\Omega_2, \Sigma_2, \mu_2)$. If f is absolutely summable or nonnegative, then

$$\int_{\Omega_1} \int_{\Omega_2} f d\mu_2 d\mu_1 = \int_{\Omega} f d\mu = \int_{\Omega_2} \int_{\Omega_1} f d\mu_1 d\mu_2.$$

2.2 Geometry of L^2[a,b]

In this section, we will develop the basic theory of the L^2 space. We start with some definitions.

Definition 2.6. The space $L^p([a,b])$ is the collection of all functions $f : [a,b] \to \mathbb{C}$ satisfying

$$\|f\|_p := \left(\int |f|^p d\mu\right)^{\frac{1}{p}} < \infty.$$

In this section, we are concerned with the case where $p = 2$.

Definition 2.7. A *Hilbert space* is an inner product space that is complete. For our purposes, we may also assume that it is separable, i.e. the space contains a countable dense subset.

Traditional examples of finite dimensional Hilbert spaces include $\mathbb{R}^n$ and $\mathbb{C}^n$. The goal of this section will be to show that L^2 is also a Hilbert space.

First, recall Schwarz's inequality, which says given an inner product space H and $f, g \in H$, we have

$$|\langle f, g\rangle| \leq \|f\|_2\|g\|_2.$$

Equality holds if and only if $f = \alpha g$, where $\alpha \in \mathbb{C}$. From this we also get the triangle inequality,

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

Note that these hold in any inner product space, including infinite dimensional ones. So, we can still use finite dimensional geometric intuition on infinite dimensional spaces.

Theorem 2.8

The space $L^2([a,b])$ is a (separable) Hilbert space.

Proof. Let $f, g \in L^2([a,b])$. Clearly it is closed under scalar multiplication. Furthermore, we have

$$\int |f+g|^2 \leq \int \left(|f|^2 + 2|f||g| + |g|^2\right) \leq \int 2\left(|f|^2 + |g|^2\right) < \infty,$$

so it is a vector space. Now, the inner product $\langle \cdot, \cdot \rangle$ given by

$$\langle f, g\rangle := \int f g^*$$

is well-defined, since

$$2|fg^*| = 2|f||g| \leq |f|^2 + |g|^2,$$

whose integral is finite. It's easy to verify the rest of the properties of an inner product.

We now move on to the meat of the theorem. It turns out that completeness doesn't even hold if we were to define L^2 in terms of the Riemann integral, so already we see an advantage of the Lebesgue integral (in fact, the "L" actually stands for Lebesgue!).

L^2 is complete. We are given a sequence $(f_n) \in L^2$ such that $\|f_n - f_m\| \to 0$ as $n, m \to 0$, and we want to show that there is a function $f \in L^2$ such that $\|f - f_n\| \to 0$. Then we can choose a subsequence so that

$$\|f_n - f_{n_j}\|^2 < 2^{-j} \quad \text{for } n \geq n_j.$$

We will show that f_{n_j} has a limit almost everywhere. Consider the "bad set" B given by

$$B = \bigcap_{i \geq 1} \bigcup_{j \geq i} B_j = B_j \text{ i.o.},$$

where $B_j = \{x : |f_{n_{j+1}}(x) - f_{n_j}(x)| \geq 2^{-j/3}\}$. Note that by Chebyshev's inequality,

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

Since this is the term of a vanishing geometric series, we have that

$$\sum_{j=1}^{\infty} \mu(B_j) < \infty.$$

Then we may apply the Borel Cantelli lemma, to deduce

$$\mu(B) = \mu(B_j \text{ i.o.}) = 0.$$

Now, note that outside of our bad set B, the sequence f_{n_j} necessarily converges pointwise, so define

$$f(x) = \begin{cases} \lim_{j\to\infty} f_{n_j}(x) & x \in [a,b] \setminus B \\ 0 & x \in B \end{cases}$$

We show that $f_n \to f$ in L^2. Since $f_{n_j} \to f$ almost everywhere, we have by Fatou's lemma that

$$\begin{aligned} \|f - f_n\|^2 &= \int_{[a,b]} \lim_{j\to\infty} |f_{n_j} - f_n|^2 \\ &\leq \liminf_{j\to\infty} \int_{[a,b]} |f_{n_j} - f_n|^2 \\ &= \liminf_{j\to\infty} \|f_{n_j} - f_n\|^2 \\ &\leq 2^{-j} \quad \text{if } n \geq n_j. \end{aligned}$$

First, this implies that $f \in L^2$, since

$$\|f\| \leq \|f - f_{n_1}\| + \|f_{n_1}\| \leq 2^{-1/2} + \|f_{n_1}\| < \infty.$$

Second, this implies that

$$\|f - f_n\| \to 0.$$

Therefore we have shown that L^2 is complete.

L^2 is separable. We are given $f \in L^2$, and we wish to find a fixed countable dense subset of L^2. Note that it suffices to show this for real valued functions, since we can easily extend to complex functions in the usual way. First, let K be the family of step functions f satisfying:

- f vanishes beyond a certain point.
- f only takes on values in $\mathbb{Q}$.
- f only jumps at a finite number of points belonging to $\mathbb{Q}$.

Note that this is a countable subset of L^2.

For any integers $i, j, k \geq 1$, define

$$\begin{aligned} f_1 &= f\mathbf{1}_{|x|\leq i} \\ f_2 &= f_1\mathbf{1}_{|f_1|\leq j} \\ f_3 &= k^{-1}\lfloor kf_2 \rfloor. \end{aligned}$$

Think of the third function f_3 as an approximation for f. Indeed note that

$$\begin{aligned} \|f - f_1\|^2 &= \int_{|x|>i} |f|^2 \\ \|f_1 - f_2\|^2 &= \int_{|f_1|>j} |f_1|^2 \leq \int_{|f|>j} |f|^2 \\ \|f_2 - f_3\|^2 &= \int_{|x|\leq i} k^{-2} \leq 2ik^{-2}. \end{aligned}$$

If we take i, j, k large enough, in that order, then we can make the above terms arbitrarily small. This implies we can make

$$\|f - f_3\| \leq \|f - f_1\| + \|f_1 - f_2\| + \|f_2 - f_3\|$$

arbitrarily small.

Now, note that $f_3 = l/k$ on the set

$$A = \{|x| \leq i\} \cap \{l/k \leq f_2 < (l+1)/k\}.$$

In particular f_3 is a simple function that is rational valued, though it's indicator functions $\mathbf{1}_A$ may have irrational jump points, and infinitely many too. But note that $\mu(A) \leq 2i$, so therefore we can approximate A by covering of intervals:

$$0 \leq \bigcup_{m=1}^{\infty} \mu(I_m) - \mu(A) \leq \sum_{m=1}^{\infty} \mu(I_m) - \mu(A) < \epsilon.$$

In particular, since the terms of the series are nonnegative, we can procure a finite union

$$A' = \bigcup_{m=1}^{N} I_m$$

which approximates A closely. Furthermore, we may modify the endpoints of each I_m to be rational with only a loss of an ϵ, so that we now have a finite union

$$A'' = \bigcup_{m=1}^{N} J_m$$

such that $\mathbf{1}_{A''} \in K$ and approximates f_3. But then $\mathbf{1}_{A''}$ also approximates f, concluding the proof of separability, and the theorem as well. □

Following up on the proof of separability, we may deduce another approximation result—this time by smooth functions.

Lemma 2.9

The subfamily containing functions $f \in C^\infty([a,b]) \cap L^2([a,b])$ that are compactly supported and infinitely differentiable is dense in $L^2([a,b])$.

Proof. Recall from the end of the proof of separability of L^2, we showed that the class K is dense in $L^2([a,b])$, where K consists of rational valued step functions which have a finitely many rational jump points and vanish far out. Then the idea is to "smooth" out the points of jump discontinuity so that the function is infinitely differentiable everywhere.

Let $f \in K$. Consider the function

$$g(x) = \begin{cases} 0 & x \leq 0, \\ e^{-\frac{1}{x} \cdot e^{\left(\frac{1}{x-1}\right)}} & 0 < x < 1, \\ 1 & x \geq 1, \end{cases}$$

which smooths out the unit step function into a C^∞ function. Then for each point of discontinuity of f, we can scale and shift g to replace the discontinuity with a smooth curve. Then we can progressively make each smoothing point steeper, so as to arbitrarily approximate f. The details are not very instructive so we have left them out. □

We now briefly set up the sister space of $L^2([a,b])$, namely $\ell^2(\mathbb{Z}^+)$. In later sections, this space will hold much importance in the context of Fourier series and transforms.

Definition 2.10. The infinite-dimensional complex sequence space $\ell^2(\mathbb{Z}^+)$ is the collection of sequences

$$z = (z_1, z_2, \ldots)$$

where each $z_i \in \mathbb{C}$, for $i \geq 1$, satsifying

$$\|z\| = \left(\sum_{n \geq 1} |z_n|^2 \right)^{1/2} < \infty.$$

The space $\ell^2(\mathbb{Z})$ of the complex functions defined on the integers $n = 0, \pm 1, \pm 2, \ldots$ is defined similarly, and is essentially identical to $\ell^2(\mathbb{Z}^+)$.

Theorem 2.11

The sequence space $\ell^2(\mathbb{Z}^+)$ is a Hilbert space under the inner product

$$\langle a, b\rangle = \sum a_n \overline{b_n}.$$

Proof. In verifying that ℓ^2 is an inner product space, the proof is much the same as for L^2. As such we will move on directly to showing completeness and separability.

ℓ^2 is complete. For ease of notation, we denote a sequence $z = (z_1, z_2, \dots)$ by its functional representation: $f(x)$ for $x \in \mathbb{Z}^+$. Then we are given a sequence of functions $f_n : \mathbb{Z}^+ \to \mathbb{C}$ such that

$$\|f_n - f_m\|^2 = \sum_{x=1}^{\infty} |f_n(x) - f_m(x)|^2 \to 0, \qquad \text{as } n, m \to \infty.$$

In particular, if $n, m \to \infty$, this implies for each $x \in \mathbb{Z}^+$,

$$|f_n(x) - f_m(x)| \to 0$$

uniformly. But since $\mathbb{C}$ is complete, there exists a limit f such that

$$f_n(x) \to f(x) \qquad \text{for each } x \in \mathbb{Z}^+.$$

Then we have, with an application of Dominated convergence, which is justified since the limit is uniform,

$$\begin{aligned}
\|f_n - f\| &= \sum_x |f_n(x) - f(x)|^2 \\
&= \sum_x \lim_{m\to\infty} |f_n(x) - f_m(x)|^2 \\
&= \lim_{m\to\infty} \sum_x |f_n(x) - f_m(x)|^2 \\
&= \lim_{m\to\infty} \|f_n - f_m\|.
\end{aligned}$$

But then if we take $n \to \infty$, we get

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

Thus $f_n \to f$ in norm. Note that this also implies $f \in \ell^2$, since for some n large enough,

$$\|f\| \le \|f - f_n\| + \|f_n\| < \epsilon + \|f_n\| < \infty.$$

Therefore, ℓ^2 is complete.

ℓ^2 is separable. Let $\epsilon > 0$ be given. Consider the countable family K of rational valued sequences. Given a sequence $z = (z_1, z_2, \dots)$, consider the rational sequence $k = (k_1, k_2, \dots) \in K$ satisfying

$$|z_i - k_i| < \sqrt{2^{-i}\epsilon}.$$

Such a sequence exists since the rational complex numbers are dense in $\mathbb{C}$. Then we have

$$\|z-k\| < \sum_{i\geq 1} |z_i - k_i|^2 < \sum_{i\geq 1} 2^{-i}\epsilon = \epsilon.$$

Since ϵ was arbitrary, this concludes the proof. □

2.3 Geometry of Generalized Fourier Series in L^2[a,b]

We begin with an important property of complex exponentials, which we will see as a component of Fourier series soon.

Lemma 2.12

The collection of functions,

$$e_n(x) = e^{2\pi inx}, \quad n \in \mathbb{Z}, \quad x \in [0,1],$$

is a orthonormal family in $L^2[0,1]$.

Proof. For $n \in \mathbb{Z}$, we have

$$\langle e_n, e_n\rangle = \int_0^1 e^{2\pi inx} e^{-2\pi inx} dx = \int_0^1 1 dx = 1.$$

For $n \neq m$, we have

$$\begin{aligned}\langle e_n, e_m\rangle &= \int_0^1 e^{2\pi inx} e^{-2\pi imx} dx \\ &= \int_0^1 e^{2\pi i(n-m)x} dx \\ &= \left.\frac{e^{2\pi i(n-m)x}}{2\pi i(n-m)}\right|_{0 \text{ to } 1} \\ &= 0.\end{aligned}$$

□

The following theorem gives us a general class of series $\sum c_n e_n$ which converges for any orthonormal family (e_n). Note that this only tells us that some target function f exists, but not necessarily that it's the one we want. Later we will see that Fourier series of L^2 functions in general will converge to what we want.

Lemma 2.13

For any orthonormal family of funtions $(e_n)_{n\geq 1}$ and scalars $c_n \in \mathbb{C}$ such that the sequence $c = (c_1, c_2, \dots) \in \ell^2(\mathbb{Z})$, we have

$$\left\| f - \sum_{n=1}^{N} c_n e_n \right\|_{L^2} \to 0 \quad \text{as } N \to \infty,$$

for some $f \in L^2$.

Proof. Let $f_n = \sum_{k=1}^{n} c_k e_k$. Since L^2 is complete, it suffices to show that (f_n) is Cauchy. Since (e_n) is orthonormal, we have

$$\begin{aligned} \|f_n - f_m\|^2 &= \left\langle \sum_{k=m+1}^{n} c_k e_k, \sum_{k=m+1}^{n} c_k e_k \right\rangle \\ &\leq \sum_{k=m+1}^{n} \langle c_k e_k, c_k e_k \rangle \\ &= \sum_{k=m+1}^{n} |c_k|^2 \to 0 \quad n, m \to 0, \end{aligned}$$

due to completeness of ℓ^2. It follows that f_n converges in L^2 to some $f \in L^2$. $\square$

Corollary 2.14 (Pythagorean Rule)

Let $f \in L^2$ and $c \in \ell^2$ be as in Lemma 2.13. If we understand $\sum_{n=1}^{\infty} c_n e_n$ to be a pointwise limit, then we have

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

Proof. By triangle inequality, we have

$$\|f\| - \left\| \sum_{n=1}^{N} c_n e_n \right\| \leq \left\| f - \sum_{n=1}^{N} c_n e_n \right\| \to 0.$$

Therefore

$$\left\| \sum_{n=1}^{N} c_n e_n \right\| \to \|f\|.$$

Now, we can square this, and use orthonormality to obtain

$$\sum_{n=1}^{N} |c_n|^2 \to \|f\|.$$

But limits are unique, and therefore

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

□

Definition 2.15. We say that a function f lies in the *span* of the family $(f_n)_{n\geq 1}$ if we can approximate f with arbitrary precision using finite linear combinations

$$\sum_{k=1}^{n} c_k f_k,$$

that is, if we can make

$$\left\| f - \sum_{k=1}^{n} c_k f_k \right\|$$

arbitrarily close to 0 with proper choice of n and coefficients $(c_1, \ldots, c_n)$. We would generally expect, as better approximations are sought, for the earlier coefficients to be adjusted. However, it turns out that for orthonormal families $(f_n) = (e_n)$, there is no need.

We also say that a family (f_n) *spans* L^2 if every $f \in L^2$ lies in the span of (f_n). Another way of saying this is to say that (f_n) is dense in L^2. An orthonormal *basis* of L^2 is an orthonormal family (e_n) which spans L^2.

Definition 2.16 (Generalized Fourier Series/Coefficients)**.** For an orthonormal family (e_n) and a function $f \in L^2$ we call $\hat{f}(k)$ the k-th *Fourier coefficient* of f, given by

$$\hat{f}(k) := \langle f, e_k \rangle = \int f e_k^*.$$

Note that since $\|e_k\|^2 = 1$, this implies that $\hat{f}(k)e_k$ is just the orthogonal projection of f onto e_k.

Theorem 2.17

Let (e_n) be a orthonormal family in L^2. Then for any $f \in L^2$ and $c_1, \ldots, c_n \in \mathbb{C}$, we have

$$\left\| f - \sum_{k=1}^{n} \hat{f}(k) e_k \right\| \leq \left\| f - \sum_{k=1}^{n} c_k e_k \right\|.$$

Furthermore, this bound is tight iff $c_k = \hat{f}(k)$ for each k.

Proof. We do the standard trick of adding and subtracting the desired term:

$$\begin{aligned}
\left\| f - \sum_{k=1}^{n} c_k e_k \right\|^2 &= \left\| f - \sum_{k=1}^{n} \hat{f}(k) e_k + \sum_{k=1}^{n} \hat{f}(k) e_k - \sum_{k=1}^{n} c_k e_k \right\|^2 \\
&= \left\| f - \sum_{k=1}^{n} \hat{f}(k) e_k \right\|^2 \\
&\quad + 2\mathrm{Re} \left\langle f - \sum_{k=1}^{n} \hat{f}(k) e_k, \sum_{k=1}^{n} (\hat{f}(k) - c_k) e_k \right\rangle \\
&\quad + \left\| \sum_{k=1}^{n} (\hat{f}(k) - c_k) e_k \right\|^2 .
\end{aligned}$$

By orthonormality, it's easy to see that the middle term vanishes. Therefore the desired inequality follows.

The last term, by Pythagoras, becomes

$$\sum_{k=1}^{n} |\hat{f}(k) - c_k|^2,$$

so we see that equality holds if and only if $\hat{f}(k) = c_k$ for each k. □

What this theorem says, geometrically, is that the Fourier coefficients are precisely the coefficients in the projection of f onto the finite dimensional space spanned by $(e_1, \ldots, e_n)$. So essentially, as we grow our function space by adding more and more e_j's, we can more closely approximate f through its projection.

Theorem 2.18 (Bessel's Inequality)

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

Proof. Given an orthonormal family (e_n) and $f \in L^2$, we know that the projection of f is orthogonal to the "altitude". Then by Pythagoras, we have

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

This then implies

$$0 \leq \left\| f - \sum_{k=1}^{n} \hat{f}(k) e_k \right\|^2 = \|f\|^2 - \sum_{k=1}^{n} |\hat{f}(k)|^2,$$

which gives us what we want. □

In the case where (e_n) is actually a basis of L^2, we actually get equality. It turns out that the complex exponentials from 2.12 actually form a basis, which we'll later show, so this will be true for classical Fourier series (not the generalized ones we defined earlier).

Theorem 2.19 (Plancherel Identity)
The orthonormal family (e_n) is a basis of L^2 iff

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

for every $f \in L^2$.

Proof. Suppose (1) holds for every $f \in L^2$. Then we have

$$\left\| \sum_{n=1}^{N} \hat{f}(n) e_n \right\|^2 = \sum_{n=1}^{N} |\hat{f}(n)|^2 \to \|f\|^2 \quad N \to \infty.$$

From this it follows that

$$\left\| f - \sum_{n=1}^{N} \hat{f}(n) e_n \right\|^2 = \|f\|^2 - \sum_{n=1}^{N} |\hat{f}(n)|^2 \to 0 \quad N \to \infty,$$

which implies (e_n) is a basis of L^2.

Conversely, suppose (e_n) is a basis of L^2. Then we can pick N large, along with complex coefficients c_n, so that

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

is small. Then by Theorem 2.17,

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

is also small. So, as $N \to \infty$, we have

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

It follows that

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

$\square$

It turns out that not all unit perpendicular families span (and therefore are bases of) L^2. One counterexample would be to take the Fourier basis in 2.12 and remove every odd indexed function. Another easy way to see this fact will be evident later, when we've shown that L^2 is isomorphic to ℓ^2. From this we can take the standard basis, where e_n is the sequence with a 1 in the n-th spot and 0's everywhere else, and then remove any single one of the e_n.

The following result gives a quick way to validate these examples.

Corollary 2.20

An orthonormal family (e_n) spans L^2 iff the only function $f \in L^2$ which is orthogonal to the whole family is $f \equiv 0$.

Proof. First suppose (e_n) is a basis of L^2, and let $f \in L^2$ be orthogonal to the whole family. Since it is a basis, we know that the Fourier series of f converges in L^2 to itself. But note

$$\hat{f}(n) = \langle f, e_n \rangle = 0 \qquad \text{for each } n,$$

so that $f = 0$.

For the converse, suppose the only function $f \in L^2$ orthogonal to the entire family is $f = 0$. Note that for all e_j,

$$\left\langle f - \sum_{n=1}^{\infty} \hat{f}(n) e_n, e_j \right\rangle = \langle f, e_j \rangle - \hat{f}(j) = 0. \tag{2}$$

Before we move on, we take a brief detour. Observe that for the first equality above we are using linearity across an infinite sum. We actually have to be a bit careful here, but it turns out that in Hilbert spaces, linearity does indeed hold for infinite sums. This has to do with the fact that inner products are continuous in each slot. In particular, consider the function

$$\varphi(x) = \langle x, y \rangle,$$

for some fixed y. Then for fixed x_0, we have by Schwarz inequality

$$\begin{aligned} |\varphi(x_0) - \varphi(x)| &= |\langle x_0, y \rangle - \langle x, y \rangle| \\ &= |\langle x_0 - x, y \rangle| \\ &\leq \|x_0 - x\| \cdot \|y\|, \end{aligned}$$

which can be made arbitrarily small by restricting x close to x_0. From this we get

$$\left\langle \sum_{n=1}^{\infty} x_n, y \right\rangle = \left\langle \lim_{N\to\infty} \sum_{n=1}^{N} x_n, y \right\rangle = \lim_{N\to\infty} \left\langle \sum_{n=1}^{N} x_n, y \right\rangle = \sum_{n=1}^{\infty} \langle x_n, y \rangle.$$

Now, from (2) we see that

$$f - \sum_{n=1}^{\infty} \hat{f}(n) e_n = 0,$$

since it's orthogonal to the entire family. Therefore, we have for all $f \in L^2$,

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

so by Theorem 1 we know that (e_n) is a basis of L^2. □

Though we haven't yet shown that $(e^{2\pi inx})$ is a basis of L^2, it's already possible to deduce that some orthornormal basis of L^2 exists without considering the functions $e^{2\pi inx}$. To see this, recall that L^2 is separable, and so there exists a countably dense subset, say (f_n). Apply Gram-Schmidt to (f_n) to obtain an orthonormal family (e_n). Since every $f \in L^2$ can be closely approximated by finite linear combinations of the (f_n), the same follows for the family (e_n), since Gram-Schmidt only uses finite linear combinations of the (f_n). Therefore, L^2 has an orthonormal basis.

Theorem 2.21 (Riesz-Fischer)
There is an isometry between L^2 and ℓ^2. That is, there is a vector space isomorphism preserving inner products.

Proof. We may work in some orthonormal basis (e_n) of L^2. Consider the map $\wedge : L^2 \to \ell^2$ given by

$$\wedge : f \mapsto \hat{f}.$$

The Plancherel identity gives us

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

so $\wedge$ preserves distances. Linearity is due to bilinearity of the inner product.

Let $f_1, f_2 \in L^2$ such that $\hat{f}_1 = \hat{f}_2$. But we know the Fourier series of f_i converge to f_i, so $f_1 = f_2$, and $\wedge$ is 1:1.

Now, for every $c = (c_1, c_2, \dots) \in \ell^2$, consider $f = \sum_{n=1}^{\infty} c_n e_n$, which we know converges in L^2 due to Lemma 2.13. Then for each $k \geq 1$,

$$\hat{f}(k) = \langle f, e_k \rangle = \left\langle \sum_{n=1}^{\infty} c_n e_n, e_k \right\rangle = c_k.$$

Thus $\wedge$ is onto.

Finally, we show $\wedge$ preserves inner products. From the Plancherel identity and linearity, we have

$$\|f_1 - f_2\|^2 = \|\hat{f}_1 - \hat{f}_2\|^2.$$

Expanding and cancelling terms gives us

$$\mathrm{Re}\langle f_1, f_2 \rangle = \mathrm{Re}\langle \hat{f}_1, \hat{f}_2 \rangle.$$

Substituting if_1 for f_1 gives us

$$\mathrm{Im}\langle f_1, f_2\rangle = \mathrm{Im}\langle \hat{f}_1, \hat{f}_2\rangle.$$

Together, this implies

$$\langle f_1, f_2\rangle = \langle \hat{f}_1, \hat{f}_2\rangle. \tag{3}$$

This concludes the proof. $\square$

Equation (3) is known as the *Parseval identity*. In some contexts the names Parseval and Plancherel are swapped. Since Parseval is just a generalization of Plancherel, we will use the two interchangeably.