

(October 27, 2016)

Introduction to Fourier series

Paul Garrett garrett@math.umn.edu <http://www.math.umn.edu/~garrett/>

[This document is http://www.math.umn.edu/~garrett/m/fun/notes.2016-17/02b_Fourier_intro.pdf]

1. Pointwise convergence
2. Fourier-Dirichlet kernel versus approximate identities
3. Fejer kernel
4. Density of trigonometric polynomials in $C^o(\mathbb{T})$
5. Exponentials form a Hilbert-space basis for $L^2(\mathbb{T})$

In his 1822 treatise on *heat*, J. Fourier espoused the wonderful idea that any function on an interval, say $[0, 1]$, could be represented as an superposition of functions $\sin 2\pi nx$ and $\cos 2\pi nx$. Equivalently, as superposition of $e^{2\pi inx}$ with $n \in \mathbb{Z}$. Since these functions are eigenfunctions for the operator d/dx , such expressions facilitated solution of differential equations.

The issue of *convergence*, which in those days could only have meant *pointwise convergence*, was recognized. The first publication proving pointwise convergence was by P. Dirichlet in 1829, although the device in the proof had appeared in an earlier manuscript of Fourier whose publication was delayed.

B. Riemann's 1854 *Habilitationschrift* concerned the representability of functions by trigonometric series.

In 1915 N. Luzin conjectured that Fourier series of functions in $L^2(\mathbb{T})$ converge almost everywhere pointwise. Decades later, in 1966, L. Carlson proved Luzin's conjecture. In 1968, R. Hunt generalized this to $L^p(\mathbb{T})$ functions for $p > 1$.

From the other side, in 1876 P. du Bois-Reymond found a *continuous* function whose Fourier series diverges at a single point. Via the *uniform boundedness theorem*, we will show later that there are continuous functions whose Fourier series diverges at any given countable collection of points. A. Kolmogorov (1923/26) gave an example of an $L^1(\mathbb{T})$ function whose Fourier series diverges pointwise *everywhere*.

The *density* of trigonometric polynomials (*finite* Fourier series) in the space of continuous function $C^o(\mathbb{T})$ can be made to follow from Weierstraß' approximation theorem. We give a somewhat different proof of density of trigonometric polynomials in $C^o(\mathbb{T})$, introducing and using the *Fejér kernel*. In 1904, L. Fejér gave an even more direct proof of the density of trigonometric polynomials in $L^2(\mathbb{T})$, in effect using an *approximate identity* made directly in terms of trigonometric polynomials. We reproduce this proof.

Urysohn's lemma implies the density of $C^o(\mathbb{T})$ in $L^2(\mathbb{T})$, so we have the *completeness* of Fourier series in $L^2(\mathbb{T})$. Thus, Fourier series of $L^2(\mathbb{T})$ functions converge to them *in the $L^2(\mathbb{T})$ topology*.

A. Zygmund's *Trigonometric Series, I, II* contains much more bibliographic and historical information.

1. Pointwise convergence

On $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with coordinates $x \rightarrow e^{2\pi ix}$, the *Fourier series* of $f \in L^2(\mathbb{T})$ is^[1]

$$f \sim \sum_{n \in \mathbb{Z}} \hat{f}(n) \cdot \psi_n \quad (\text{with } \psi_n(x) = e^{2\pi inx} \text{ and } \hat{f}(n) = \int_{\mathbb{T}} f \cdot \bar{\psi}_n)$$

[1] Temporarily, to integrate a function F on the quotient $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, it suffices to let the quotient map be $q : \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$, and integrate $F \circ q$ on a convenient set of representatives for the quotient, such as $[0, 1]$ or any other interval $[x_o, x_o + 1]$. In fact, we will suppress reference to the quotient map and identify functions on \mathbb{T} with periodic functions on \mathbb{R} . A more systematic approach to integration on quotients, that does not require determination of nice sets of representatives, will be discussed later.

We do not write *equality* of the function and its Fourier series, since the question of possible *senses* of equality is significant. After all, the right-hand side is an infinite sum, possibly *numerical*, but also possibly of *functions*, and the latter offers several potential interpretations. It is completely natural to ask for *pointwise* convergence of a Fourier series, and, implicitly, convergence to the function of which it is the Fourier series. We address this first.

We can prove pointwise convergence even before proving that the exponentials give an orthonormal *basis* for $L^2[0, 1]$. The hypotheses of the convergence claim below are not optimal, but are sufficient for some purposes, and are tangible.

We also need:

[1.1] **Claim:** (*Riemann-Lebesgue*) For $f \in L^2(\mathbb{T})$, the Fourier coefficients $\widehat{f}(n)$ of f go to 0.

Proof: The $L^2[0, 1]$ norm of $\psi_n(x) = e^{2\pi i n x}$ is 1. *Bessel's inequality*

$$\|f\|_{L^2}^2 \geq \sum_n \left| \langle f, \psi_n \rangle \right|^2$$

from abstract Hilbert-space theory applies to an orthonormal *set*, whether or not it is an orthonormal *basis*. Thus, the sum on the right converges, so by Cauchy's criterion the summands go to 0. ///

A function f on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is (*finitely*) *piecewise* C^o when there are finitely many real numbers $a_0 \leq a_1 \leq \dots \leq a_{n-1} \leq a_n = a_0 + 1$ and C^0 functions f_i on $[a_i, a_{i+1}]$ such that

$$f_i(x) = f(x) \quad \text{on } [a_i, a_{i+1}] \quad (\text{except possibly at the endpoints})$$

Thus, while $f_i(a_{i+1})$ may differ from $f_{i+1}(a_{i+1})$, and $f(a_{i+1})$ may be different from both of these, the function f is continuous in the interiors of the intervals, and behaves well *near* the endpoints, if not *at* the endpoints.

Write

$$\langle f, F \rangle = \int_{\mathbb{T}} f \cdot \overline{F} = \int_0^1 f(x) \overline{F(x)} dx$$

and

$$\widehat{f}(n) = \langle f, \psi_n \rangle = \int_{\mathbb{T}} f \cdot \overline{\psi_n} = \int_0^1 f(x) \overline{\psi_n(x)} dx$$

[1.2] **Claim:** Let f be finitely piecewise C^o on \mathbb{T} . Let x_o be a point at which f has both *left and right derivatives* (even if they do not agree), and is *continuous*. Then the Fourier series of f evaluated at x_o converges to $f(x_o)$. That is,

$$f(x_o) = \sum_{n \in \mathbb{Z}} \widehat{f}(n) \psi_n(x_o) \quad (\text{a convergent sum})$$

Proof: First, make reductions to unclutter the notation. By considering $f(x) - f(x_o)$, and observing that *constants* are represented pointwise by their Fourier expansions, we can assume that $f(x_o) = 0$. The Fourier coefficients of *translates* of a function f are expressible in terms of the Fourier coefficients of f itself, using the periodicity of f as a function on \mathbb{R} :

$$\begin{aligned} \int_0^1 f(x+x_o) \overline{\psi_n(x)} dx &= \int_{x_o}^{1+x_o} f(x) \overline{\psi_n(x-x_o)} dx = \int_{x_o}^1 f(x) \overline{\psi_n(x-x_o)} dx + \int_1^{1+x_o} f(x) \overline{\psi_n(x-x_o)} dx \\ &= \int_{x_o}^1 f(x) \overline{\psi_n(x-x_o)} dx + \int_0^{x_o} f(x) \overline{\psi_n(x-x_o)} dx = \int_0^1 f(x) \overline{\psi_n(x-x_o)} dx \\ &= \psi_n(x_o) \int_0^1 f(x) \overline{\psi_n(x)} dx = \psi_n(x_o) \cdot \widehat{f}(n) \end{aligned}$$

The left-hand side is the n^{th} Fourier coefficient of the translate $x \rightarrow f(x + x_o)$, that is, the n^{th} Fourier term of $x \rightarrow f(x + x_o)$ evaluated at 0, while the right-hand side is 2π times the n^{th} Fourier term of $f(x)$ evaluated at x_o . Thus, we can simplify further by taking $x_o = 0$, without loss of generality. [2]

A partial sum of the Fourier expansion evaluated at 0 is

$$\begin{aligned} \sum_{-M \leq n < N} \int_0^2 f(x) \bar{\psi}_n(x) dx &= \int_0^1 f(x) \sum_{-M \leq n < N} \bar{\psi}_n(x) dx \\ &= \int_0^1 f(x) \frac{\bar{\psi}_N(x) - \bar{\psi}_{-M}(x)}{\psi_{-1}(x) - 1} dx \end{aligned}$$

by summing the geometric series. This is

$$\int_0^1 \frac{f(x)}{\psi_{-1}(x) - 1} (\bar{\psi}_N(x) - \bar{\psi}_{-M}(x)) dx = \left\langle \frac{f}{\psi_{-1} - 1}, \psi_N \right\rangle - \left\langle \frac{f}{\psi_{-1} - 1}, \psi_{-M} \right\rangle$$

The latter two terms are Fourier coefficients of $f/(\psi_{-1} - 1)$, so go to 0 by the Riemann-Lebesgue lemma for $f(x)/(\psi_{-1}(x) - 1)$ in $L^2(\mathbb{T})$. Since $x_o = 0$ and $f(x_o) = 0$

$$\frac{f(x)}{\psi_{-1}(x) - 1} = \frac{f(x)}{x} \cdot \frac{x}{\psi_{-1}(x) - 1} = \frac{f(x) - f(x_o)}{x - x_o} \cdot \frac{x - x_o}{e^{-2\pi i x} - e^{-2\pi i x_o}}$$

The existence of left and right derivatives of f at $x_o = 0$ is exactly the hypothesis that this expression has left and right limits at x_o , even if they do not agree.

At all other points the division by $\psi_{-1}(x) - 1$ does not disturb the continuity. Thus, $f/(\psi_{-1} - 1)$ is still at least *continuous* on each interval $[a_i, a_{i+1}]$ on which f was essentially a C^0 function. Therefore, $f/(\psi_{-1} - 1)$ is continuous on a finite set of closed (finite) intervals, so bounded on each one. Thus, $f/(\psi_{-1} - 1)$ is indeed L^2 , and we can invoke Riemann-Lebesgue to see that the integral goes to $0 = f(x_o)$. ///

[1.3] Corollary: The Fourier series of $f \in C^1(\mathbb{T})$ converges pointwise to f everywhere. ///

[1.4] Remark: Pointwise convergence does not give L^2 convergence, and we have *not* yet proven that the exponentials are an orthonormal *basis* for the Hilbert space $L^2(\mathbb{T})$. The pointwise result just proven is suggestive, but not decisive.

[2] The rearrangement of the integral would have been simpler if we integrated directly on \mathbb{T} , a *group*, rather than on a set of representatives in \mathbb{R} which had to be rearranged. That is, the change of variables that replaces x by $x - x_o$ is an automorphism of \mathbb{T} , and the measure is invariant under translation, so for F on \mathbb{T} we can write simply

$$\int_{\mathbb{T}} F(x) dx = \int_{\mathbb{T}} F(x - x_o) dx$$

2. Fourier-Dirichlet kernel versus approximate identities

Under suitable hypotheses on f , in the above proof of pointwise convergence, rewriting a little, we have

$$\begin{aligned} f(x) &= \lim_N \int_{\mathbb{T}} f(x + \xi) \cdot \left(\sum_{-N \leq n \leq N} \bar{\psi}_n(\xi) \right) d\xi = \lim_N \int_{\mathbb{T}} f(x + \xi) \cdot \frac{e^{2\pi i N \xi} - e^{-2\pi i (N+1)\xi}}{1 - e^{-2\pi i \xi}} d\xi \\ &= \lim_N \int_{\mathbb{T}} f(x + \xi) \cdot \frac{e^{2\pi i (N + \frac{1}{2})\xi} - e^{-2\pi i (N + \frac{1}{2})\xi}}{e^{\pi i \xi} - e^{-\pi i \xi}} d\xi = \lim_N \int_{\mathbb{T}} f(x + \xi) \cdot \frac{\sin(2\pi(N + \frac{1}{2})\xi)}{\sin(\pi\xi)} d\xi \end{aligned}$$

The sequence of functions

$$K_N(\xi) = \frac{\sin(2\pi(N + \frac{1}{2})\xi)}{\sin(\pi\xi)} d\xi$$

are often called the *Dirichlet kernel(s)*, although these functions did appear earlier in work of Fourier himself, whose publication was delayed.

Unlike the *Fejér kernel* in the following section, the Fourier-Dirichlet kernel does *not* have properties that would make it an *approximate identity*. An *approximate identity* on \mathbb{T} is a sequence $\{\varphi_1, \varphi_2, \dots\}$ of continuous functions such that

$$\int_{\mathbb{T}} \varphi_n = 1 \quad (\text{for all } n)$$

and such that the *masses* bunch up near $1 \in \mathbb{T}$, in the sense that for every neighborhood U of 1 in \mathbb{T} ,

$$\lim_n \int_U \varphi_n \rightarrow 1$$

The virtue of an approximate identity, not possessed by the Fourier-Dirichlet kernel, is

[2.1] **Claim:** For an approximate identity $\{\varphi_n\}$ on \mathbb{T} and for $f \in C^o(\mathbb{T})$,

$$\lim_n \int_{\mathbb{T}} f(x + \xi) \varphi_n(\xi) d\xi = f(x) \quad (\text{uniformly in } x \in \mathbb{T})$$

Proof: By the uniform continuity of f on compact \mathbb{T} , given $\varepsilon > 0$, there is $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all x, y with $|x - y| < \delta$, in terms of the parametrization $\mathbb{R} \rightarrow \mathbb{T}$.

Let N be the image in \mathbb{T} of $(-\delta, \delta) \subset \mathbb{R}$. Invoking the approximate identity property, let n_o be large enough so that $|\int_N \varphi_n - 1| < \varepsilon$ for all $n \geq n_o$. Since $\int_{\mathbb{T}} \varphi_n = 1$ and $\varphi_n(\xi) \geq 0$ for all ξ , this also implies $|\int_{\mathbb{T}-N} \varphi_n| < \varepsilon$. Then

$$\int_{\mathbb{T}} f(x + \xi) \varphi_n(\xi) d\xi = \int_N f(x + \xi) \varphi_n(\xi) d\xi + \int_{\mathbb{T}-N} f(x + \xi) \varphi_n(\xi) d\xi$$

The first integral, over U , is

$$\int_U f(x) \varphi_n(\xi) d\xi + \int_U (f(x + \xi) - f(x)) \varphi_n(\xi) d\xi = f(x) \cdot \int_U \varphi_n(\xi) d\xi + \int_U (f(x + \xi) - f(x)) \varphi_n(\xi) d\xi$$

As $n \rightarrow +\infty$, the first summand goes to $f(x) \cdot 1$ uniformly in x . The second summand is small:

$$\left| \int_U (f(x + \xi) - f(x)) \varphi_n(\xi) d\xi \right| = \int_U |f(x + \xi) - f(x)| \cdot \varphi_n(\xi) d\xi < \int_U \varepsilon \cdot \varphi_n(\xi) d\xi \leq \varepsilon \int_{\mathbb{T}} \varphi_n(\xi) d\xi = \varepsilon$$

Similarly, the integral over $\mathbb{T} - U$ is small, uniformly in x :

$$\left| \int_{\mathbb{T}-U} f(x + \xi) \varphi_n(\xi) d\xi \right| \leq \sup_{y \in \mathbb{T}} |f(y)| \cdot \int_{\mathbb{T}-U} \varphi_n(\xi) d\xi < \sup_{y \in \mathbb{T}} |f(y)| \cdot \varepsilon$$

giving the assertion of the claim. ///

3. Fejer kernel

We give Fejér's approximate identity consisting of trigonometric polynomials (finite Fourier series), whose property rearranges to prove sup-norm convergence of a sequence of trigonometric polynomials to given $f \in C^o(\mathbb{T})$.

However, the trigonometric polynomials converging uniformly pointwise to $f \in C^o(\mathbb{T})$ are *not* the finite partial sums of the Fourier series of f , but, rather the *Cesaro-summed* version of these partial sums. That is, given a sequence b_1, b_2, \dots , the Cesaro-summed sequence is

$$s_1 = \frac{b_1}{1} \quad s_2 = \frac{b_1 + b_2}{2} \quad s_3 = \frac{b_1 + b_2 + b_3}{3} \quad s_4 = \frac{b_1 + b_2 + b_3 + b_4}{4} \quad \dots$$

On one hand, if the original sequence converges, then the Cesaro-summed sequence also converges, with the same limit. On the other hand, the Cesaro-summed sequence may converge though the original does not.

As it happens, Cesaro-summing the sequence of Fourier-Dirichlet kernels $K_N(x) = \sin(2\pi(N + \frac{1}{2})x) / \sin(\pi x)$ produces an approximate identity: the Fejér kernel is

$$F_n(x) = \frac{1}{n} \sum_{j=1}^n K_{j-1}(x) = \frac{1}{n} \sum_{j=0}^{n-1} \left(\sum_{-j \leq \ell \leq j} e^{2\pi \ell x} \right) = \sum_{-(n-1) \leq \ell \leq n-1} \frac{n - |\ell|}{n} e^{2\pi \ell x} = \sum_{-n \leq \ell \leq n} \frac{n - |\ell|}{n} e^{2\pi \ell x}$$

Visibly, $F_n(x)$ is a finite Fourier series. As with the Fourier-Dirichlet kernel, we can sum geometric series to simplify:

[3.1] Claim:

$$F_n(x) = \frac{1}{n} \cdot \frac{1 - \cos 2\pi n x}{1 - \cos 2\pi x}$$

In particular, $F_n(x) \geq 0$ for all x .

Proof: From

$$K_n(x) = \frac{\sin(2\pi(n + \frac{1}{2})x)}{\sin(\pi x)}$$

computing directly,

$$\begin{aligned} \sum_{j=1}^n \frac{\sin(2\pi(j - \frac{1}{2})x)}{\sin(\pi x)} &= \frac{1}{2i \sin \pi x} \sum_{j=1}^n \left(e^{2\pi i(j - \frac{1}{2})x} - e^{-2\pi i(j - \frac{1}{2})x} \right) \\ &= \frac{1}{2i \sin \pi x} \left(\frac{e^{\pi i x} - e^{2\pi i(n + \frac{1}{2})x}}{1 - e^{2\pi i x}} - \frac{e^{-\pi i x} - e^{-2\pi i(n + \frac{1}{2})x}}{1 - e^{-2\pi i x}} \right) \\ &= \frac{1}{2i \sin \pi x} \left(\frac{1 - e^{2\pi i n x}}{e^{-\pi i x} - e^{\pi i x}} - \frac{1 - e^{-2\pi i n x}}{e^{\pi i x} - e^{-\pi i x}} \right) = \frac{1}{2i \sin \pi x} \frac{e^{2\pi i n x} - 2 + e^{-2\pi i n x}}{e^{\pi i x} - e^{-\pi i x}} \\ &= \frac{1}{2i \sin \pi x} \frac{2(\cos 2\pi n x - 1)}{2i \sin \pi x} = \frac{1 - \cos 2\pi n x}{2(\sin \pi x)^2} = \frac{1 - \cos 2\pi n x}{1 - \cos 2\pi x} \end{aligned}$$

as asserted. ///

We check the other properties for $\{F_n\}$ to be an approximate identity:

[3.2] Claim: $\int_{\mathbb{T}} F_n(x) dx = 1$, and

$$\int_{|x| \leq \frac{1}{\sqrt{n}}} F_n(x) dx \longrightarrow 0 \quad (\text{as } n \rightarrow \infty)$$

Proof: First,

$$\int_{\mathbb{T}} F_n(x) dx = \sum_{-n \leq \ell \leq n} \frac{n - |\ell|}{n} \int_{\mathbb{T}} e^{2\pi i \ell x} dx = \sum_{-n \leq \ell \leq n} \frac{n - |\ell|}{n} \begin{cases} 0 & (\text{for } \ell \neq 0) \\ 1 & (\text{for } \ell = 0) \end{cases} = 1$$

To show that the masses bunch up at 0, note that on the interval $[-\frac{1}{2}, \frac{1}{2}]$,

$$F_n(x) = \frac{1}{n} \cdot \frac{1 - \cos 2\pi n x}{1 - \cos 2\pi x} = \frac{1}{n} \cdot \frac{(1 - \cos 2\pi n x)/x^2}{(1 - \cos 2\pi x)/x^2}$$

The denominator $(1 - \cos 2\pi x)/x^2$ is non-vanishing and continuous on that interval, so is uniformly bounded away from 0. Thus, it suffices to show that the integral of $(1 - \cos 2\pi n x)/nx^2$ on $[-\frac{1}{2}, \frac{1}{2}]$ outside $[-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}]$ goes to 0. Indeed,

$$\frac{1}{n} \int_{\frac{1}{\sqrt{n}}}^1 \frac{1 - \cos 2\pi n x}{x^2} dx = n \cdot \int_{\frac{1}{\sqrt{n}}}^1 \frac{1 - \cos 2\pi n x}{(nx)^2} dx = \int_{\sqrt{n}}^n \frac{1 - \cos 2\pi x}{x^2} dx$$

by replacing x by x/n . This is dominated by

$$\int_{\sqrt{n}}^{\infty} \frac{dx}{x^2} = \frac{1}{\sqrt{n}} \longrightarrow 0$$

This proves that $F_n(x)$ forms an approximate identity. ///

4. Completeness of Fourier series in $L^2(\mathbb{T})$

Using the approximate identity property of the Fejér kernels F_n , we can prove

[4.1] Corollary: The vector space of finite trigonometric polynomials is dense in $C^o(\mathbb{T})$, and, hence, in $L^2(\mathbb{T})$.

Proof: On one hand, from the discussion of approximate identities,

$$\int_{\mathbb{T}} F_n(\xi) f(x + \xi) d\xi \longrightarrow f(x) \quad (\text{in sup-norm})$$

On the other hand, by rearranging and changing variables [3]

[3] As earlier, if we imagine we are integrating on an interval, then a change of variables entails breaking the interval into two pieces and rearranging. This necessity is avoided if we know how to integrate on groups \mathbb{T} , whether or not expressible as quotients \mathbb{R}/\mathbb{Z} .

$$\begin{aligned}\int_{\mathbb{T}} F_n(\xi) f(x + \xi) d\xi &= \sum_{|\ell| \leq n} \frac{n - |\ell|}{n} \int_{\mathbb{T}} e^{2\pi i n \xi} f(x + \xi) d\xi = \sum_{|\ell| \leq n} \frac{n - |\ell|}{n} \int_T e^{2\pi i n(\xi - x)} f(\xi) d\xi \\ &= \sum_{|\ell| \leq n} \left(\frac{n - |\ell|}{n} \int_T e^{2\pi i n \xi} f(\xi) d\xi \right) \cdot e^{-2\pi i n x}\end{aligned}$$

That is, these trigonometric polynomials approach $f \in C^0(\mathbb{T})$ in sup-norm. ///

5. Exponentials form a Hilbert-space basis for $L^2(\mathbb{T})$

[5.1] Corollary: The exponentials $x \rightarrow e^{2\pi i n x}$ for $n \in \mathbb{Z}$ are an orthonormal basis, that is, a Hilbert space basis, for $L^2(\mathbb{T})$.

Proof: That they are mutually orthogonal, and have L^2 -norms all 1, is a direct computation. The density of trigonometric polynomials in $L^2(\mathbb{T})$ is the assertion that the vector space of finite linear combinations of these exponentials is *dense* in $L^2(\mathbb{T})$. That is, there are no (non-zero) vectors orthogonal to all exponentials. ///

[5.2] Corollary: The Fourier series $\sum_n \langle f, \psi_n \rangle \cdot \psi_n$ of $f \in L^2(\mathbb{T})$ converges to f in the L^2 topology. ///
