III. *An Introduction to Fourier Series*

Joel H. Shapiro

January 15, 2018

These notes introduce the notion of Fourier series for Lebesgue-integrable functions on the interval $[-\pi, \pi]$. The goal is to prove Fejér's Theorem: *The arithmetic means of the symmetric partial sums of the Fourier series of such a function converges to the function in the L^1-norm*. In a subsequent revision we'll show that these arithmetic means converge to the function almost everywhere. This stands in sharp contrast to Kolmogorov's 1923 result [4]: *There exist functions in $L^1([-\pi, \pi])$, the symmetric partial sums of whose Fourier series* diverge *a.e.*

0 *Complex Exponentials*

Recall the *complex exponential*

$$e^{ix} := \cos x + i \sin x \qquad (x \in \mathbb{R}).$$

Thanks to the Pythagorean Theorem, $|e^{ix}| = 1$ for each real x, and thanks to the addition laws for sine and cosine:

$$e^{i(x+y)} = e^{ix}e^{iy} \qquad (x, y \in \mathbb{R}).$$

In particular:

$$e^{ix}e^{-ix} = 1 \qquad (x \in \mathbb{R}),$$

$\mathbb{R}$ denotes the real line, $\mathbb{Z}$ the integers.

and

$$(e^{ix})^n = e^{inx} \qquad (x \in \mathbb{R},\ n \in \mathbb{Z}).$$

One also checks easily that for $n \in \mathbb{Z}$:

$$(1) \qquad \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\,dx = \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ if } n \neq 0. \end{cases}$$

The integral of a complex-valued function is the integral of its real part plus i times that of its imaginary part.

1 *Trigonometric Series*

A *trigonometric series* is a "formal" sum expressed as either an "infinite linear combination" of complex exponentials:

Here "formal" means we make about the convergence—if there is any—of the infinite series.

$$(2) \qquad \sum_{n\in\mathbb{Z}} c_n e^{inx}$$

with complex coefficients c_n, or an infinite linear combination of sines and cosines:

$$(3) \qquad a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

with real coefficients a_n and b_n. It's an easy exercise to check that the "real part" (resp. "imaginary part") of the series (2) has the form (3).

A *trigonometric polynomial* is a finite linear combination of complex exponentials (complex case), or of sines and cosines (real case). To say the trigonometric series (2) *converges* (at $x \in \mathbb{R}$, pointwise on $\mathbb{R}$, uniformly ...) means that its sequence of *symmetric partial sums* $(\sum_{|n|\leq N} c_n e^{inx})_0^\infty$ (each of which is a trigonometric polynomial) converges.

Trigonometric series arose in the eighteenth and nineteenth centuries in the course of solving problems of wave motion and heat flow. For this it was important to determine, for a function f represented—in some appropriate sense—by a trigonometric series, a formula for the coefficients of its trigonometric series expansion.

The freely downloadable lecture slides [1] give a beautiful panoramic history of Fourier analysis, both in its early and its more modern stages.

In this regard, suppose f is a trigonometric polynomial:

$$f(x) = \sum_{|n|\leq N} c_n e^{inx}.$$

Then for $k \in \mathbb{Z}$:

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{|k|\leq N} c_n e^{ikx}\right)e^{-inx}\,dx \\
&= \sum_{|k|\leq N} c_n \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-inx}\,dx\right) \\
&= \sum_{|k|\leq N} c_n \underbrace{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-n)x}\,dx\right)}_{=0 \text{ if } k\neq n\,,=1 \text{ if } k=n}
\end{aligned}$$

Thus (now swapping the roles of indices k and n):

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx \qquad (n \in \mathbb{Z}).$$

The argument works as well for functions f represented by trigonometric series (2) with $\sum_{n\in\mathbb{Z}} |c_n| < \infty$. Indeed: since $|e^{inx}| = 1$ for each index n and each real x, such a series converges uniformly on the real line by the Weierstrass M-test.[1] Consequently its sum is continuous on $\mathbb{R}$, and the interchange of sum and integral that established the previous result still works. In summary:

[1] See, e.g., [5], Theorem 7.10, page 148.

Proposition 1.1. *If $(c_n : n \in \mathbb{Z})$ is an absolutely summable sequence of complex numbers, then the equation*

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e^{inx}$$

defines a uniformly continuous function on the real line, and

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx \qquad (n \in \mathbb{Z}). \tag{4}$$

2 Fourier Series

The largest class of functions to which to which we might be able to extend some version of Proposition 1.1 is $L^1 = L^1([-\pi, \pi])$, which—for this lecture—is the space of all (a.e. equivalence classes of) complex-valued, Lebesgue measurable, 2π-periodic functions on the real line, endowed with the norm

$$\|f\|_1 := \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|\, dx < \infty$$

(which makes L^1 into a Banach space). At the very least, Proposition 1.1 suggests a natural trigonometric series to associate with f, namely the formal series:

$$\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx} \tag{5}$$

where

$$\hat{f}(n) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\, dx \qquad (n \in \mathbb{N}). \tag{6}$$

Some terminology. For $f \in L^1$:

- The coefficient $\hat{f}(n)$ is the n-th *Fourier coefficient* of f.
- the series (5) is the *Fourier series* of f.
- The (doubly infinite) sequence $\hat{f} = (\hat{f}(n))_{n\in\mathbb{Z}}$ is the *Fourier transform of* f.

The map $\mathscr{F}\colon f \to \hat{f}$, is the *Fourier transform*; it's a linear transformation taking L^1 into the vector space of (two-sided) complex sequences.

Some questions ... suggested by what we've done so far.

(Q1) Does the Fourier series of $f \in L^1$ converge "in some natural way" to f?

(Q2) Is the Fourier transform one-to-one? What is its range?

It's easy to see that the Fourier transform of each L^1-function is a bounded sequence. In fact,

$$|\hat{f}(n)| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\, dx\right| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|\,|\underbrace{e^{-inx}}_{\equiv 1}|\, dx = \|f\|_1 .$$

Proving this "crashing absolute values through the integral" inequality, while routine for real-valued integrands, is a little bit tricky in the complex case; see, e.g., [5], page 325.

Thus, $\mathscr{F}$ actually maps L^1 into the space $\ell^\infty(\mathbb{Z})$ of *bounded* two-sided complex sequences, and the inequality just proves that

$$\|\mathscr{F}(f)\|_\infty \le \|f\|_1 \qquad (f \in L^1), \tag{7}$$

where $\|\cdot\|_\infty$ is the supremum norm on $\ell^\infty(\mathbb{Z})$. In the language of normed linear spaces: *The Fourier transform maps* L^1 contractively *into* $\ell^\infty(\mathbb{Z})$.

This suggests rephrasing the latter part of question Q2 above as:

(Q3) Does $\mathscr{F}$ map L^1 *onto* $\ell^\infty(\mathbb{Z})$? If not, what is $\mathscr{F}(L^1)$?

In fact, $\mathscr{F}(L^1)$ does *not* exhaust all of $\ell^\infty(\mathbb{Z})$. According to the Riemann-Lebesgue Lemma (Corollary 8.3 of §8 below), $\hat{f} \in c_0(\mathbb{Z})$ (those two-sided sequences (a_n) with limit zero as $n \to \infty$). Thus $\mathscr{F}(L^1) \neq \ell^\infty(\mathbb{Z})$. Is $\mathscr{F}(L^1) = c_0(\mathbb{Z})$? No again! For this see, e.g., [6], Thm. 5.15, page 104.

As for the first part of (Q2) the answer is "Yes, $\mathscr{F}$ is one-to-one." However the argument, which we'll give in §7, is not trivial—it depends on finding a positive answer to (Q1).

3 *Convergence for Fourier Series?*

For $f \in L^1$ it's traditional to write

$$f \sim \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$$

where the symbol "$\sim$" expresses the fact that we haven't yet described just how the function is represented by its Fourier series.

Some history. We've already seen (Proposition 1.1) that "$\sim$" can be replaced by "$=$" if f is a trigonometric polynomial, or more generally if $\hat{f}$ is an absolutely summable (doubly infinite) sequence. Work of Dirichlet in the early 1800's, completed by Jordan about fifty years later, showed that if f is continuous and 2π-periodic on the real line, and of *bounded variation* on $[-\pi, \pi]$, then the Fourier series of f converges to f at each point of $\mathbb{R}$. However, soon afterward, the Pierre du Bois Reymond showed that one could not omit the additional hypothesis of bounded variation.[2] Motivated by the Lebesgue theory of measure and integration, in 1920 the Russian mathematician Nikolai Lusin asked if, nevertheless, such a continuous function must be the sum of its Fourier series at *almost every* point of the real line.

[2] See e.g., Rudin [6], §5.11–5.13, pp. 100-103 for the modern treatment of this "catastrophe."

Regarding Lusin's question, another Russian, Andrei Kolmogorov (born April 1903) showed in 1923 (!) that there exist functions $f \in L^1$ whose Fourier series *diverge* at almost every point of $\mathbb{R}$,[3] and a few years later he produced an L^1-function whose Fourier series diverged at *every* point of $\mathbb{R}$. Nevertheless, Lusin's question remained unanswered until 1966, when the Swedish mathematician Lennart Carleson proved something a lot stronger:

[3] Fundamenta Math. 4, pp. 324–328.

Carleson's Theorem.[4] *If* $f \in L^2$, *then* $f(x) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{inx}$ *for almost-every* $x \in [-\pi, \pi]$.

[4] Acta Math. 116 (1966) 135–157.

4 *The partial sums*

In this section we'll fix $f \in L^1$, N a non-negative integer, and x a real number. We seek a closed-form expression for the N-th (symmetric) partial sum of the Fourier series of f:

$$s_N f(x) := \sum_{|n| \le N} \hat{f}(n) e^{inx} . \tag{8}$$

There's only one course of action: substitute the definition (6) for $\hat{f}(n)$ into the (8) and see what happens.

$$\begin{aligned}(s_N f)(x) &= \sum_{|n|\le N} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int}\, dx \right) e^{inx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{|n|\le N} e^{-int} e^{inx} \right) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{|n|\le N} e^{in(x-t)} \right) dt\end{aligned}$$

Thus

$$(s_N f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x-t) f(t)\, dt , \tag{9}$$

where

$$D_N(x) = \sum_{|n|\le N} e^{inx} . \tag{10}$$

D_N is called the N-th *Dirichlet kernel*.

Consequently the problem of finding a closed-form expression for $s_N f(x)$ reduces to that of finding such an expression for D_N. The trick is to use note that $e^{inx} = (e^{ix})^n$ and use the addition formula for the complex exponential to reduce the problem to summation of a geometric-series partial sum. More precisely:

$$D_N(x) = e^{-iNx} \sum_{n=0}^{2N} (e^{ix})^n = e^{-iNx} \frac{1 - (e^{ix})^{2N+1}}{1 - e^{ix}}$$

$$= \underbrace{e^{-iNx} \cdot \frac{e^{i\frac{2N+1}{2}x}}{e^{ix/2}}}_{\equiv 1} \cdot \underbrace{\frac{e^{-i\frac{2N+1}{2}x} - e^{i\frac{2N+1}{2}x}}{e^{-ix/2} - e^{ix/2}}}_{= \frac{\sin[(2N+1)x/2]}{\sin(x/2)}}$$

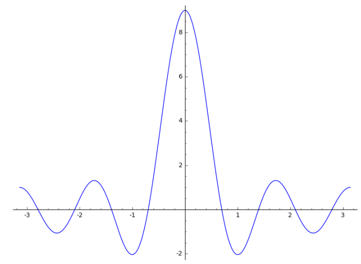
Summarizing:

Proposition 4.1. *For $f \in L^1$, $x \in \mathbb{R}$, and N a non-negative integer:*

$$(s_N f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x-t) f(t)\, dt\,, \tag{11}$$

where

$$D_N(x) = \frac{\sin\left((2N+1)\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)}\,.$$



The Dirichlet kernel D_4

5 *Convolution integrals*

The integral that shows up in the representation (11) of Fourier-series partial sums is a special *convolution integral*. More generally, if f and g belong to L^1 then their *convolution* is defined by:

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) g(t)\, dt \qquad (x \in \mathbb{R}). \tag{12}$$

While the integral on the right-hand side of (12) is clearly defined if, for example, one of the functions involved is bounded (as is the case in eqn. (11) above), it's not clear that this is generally true. Fortunately it is. To see why, let's first suppose that our L^1-functions f and g take only non-negative values. Then the right-hand side of (12) is defined, although its value may be $+\infty$. In any case, we can integrate it and use Fubini's Theorem to interchange the order of integration.[5] For brevity we'll use the notation "$\int$" to signify "$\int_{-\pi}^{\pi}$".

[5] Strictly speaking: we should check that the integrand satisfies the measurability hypotheses of Fubini's theorem. Rule of Thumb: The "Fubini-theorem hypotheses are always satisfied."

$$\begin{aligned}
\int (f*g)(x)\, dx &= \int \left(\int f(x-t) g(t)\, dt \right) dx \\
&= \int \left(\int f(x-t) g(t)\, dx \right) dt \\
&= \int \left(\int f(x-t)\, dx \right) g(t)\, dt \\
&= \int \left(\int f(x)\, dx \right) g(t)\, dt \\
&= \int f(x)\, dx \int g(t)\, dt\,.
\end{aligned}$$

The equality $\int f(x-t)\, dx = \int f(x)\, dx$ used in the last line is a consequence of the 2π-periodicity of f; see Proposition 6.1 below.

Thus, if f and g are non-negative functions belonging to L^1, then $f * g$ belongs to L^1 with $\|f * g\|_1 = \|f\|_1 \cdot \|g\|_1$. In particular, the integral defining $(f * g)(x)$ is finite for a.e. $x \in \mathbb{R}$.

If f and g are arbitrary L^1-functions, the argument above shows that the convolution of their respective absolute-value functions is finite a.e., i.e., that the function $t \to f(x-t)g(t)$ belongs to L^1 for a.e. $x \in \mathbb{R}$. Thus it makes sense to form the convolution integral $(f * g)(x)$ for a.e. x. Moreover, for such x:

$$|(f*g)(x)| = \left| \int f(x-t) g(t)\, dt \right| \leq \int |f(x-t)| \cdot |g(t)|\, dt = (|f| * |g|)(x)\,.$$

Thus we've established, modulo proving that $\int f(x-t)\,dx = \int f(x)\,dx$ in the calculation above:

Proposition 5.1. *If f and g belong to L^1 then so does $f * g$, and*

$$\|f * g\|_1 \le \|f\|_1 \cdot \|g\|_1 .$$

6 *Translation-invariance of integrals*

Here we'll tie up that loose end from the proof of Proposition 5.1 by proving:

Proposition 6.1. *If $f \in L^1$ then $\int_{-\pi}^{\pi} f(x-t)\,dx = \int_{-\pi}^{\pi} f(x)\,dx$ for every $t \in \mathbb{R}$.*

For the proof, note that since each $f \in L^1$ is 2π-periodic on $\mathbb{R}$, the same is true, for each $t \in \mathbb{R}$, of f_t, its of its "translate by t," defined as:

$$f_t(x) = f(x-t) \qquad (x \in \mathbb{R}). \tag{13}$$

By a simple change of variable,

$$\int_{-\pi}^{\pi} f(x-t)\,dx = \int_{-\pi-t}^{\pi-t} f(y)\,dy\,,$$

so our problem reduces to one of showing that each $f \in L^1$ has the same integral over any interval of length 2π. More generally:

Lemma 6.2. *Suppose f is a real or complex-valued function that is Lebesgue measurable on $\mathbb{R}$ and periodic with period $T > 0$. Then $\int_I f(x)\,dx = \int_J f(x)\,dx$ for any real intervals I and J of length T.*

Proof. By a simple change of variable it's enough to do this for $T = 1$. We may take $I = [0,1]$. Write $J = [a, a+1]$ for some $a \in \mathbb{R}$. Then $a = b + n$ for some integer n and some $b \in [0,1)$. Then

$$\int_a^{a+1} f(x)\,dx = \int_{b+n}^{b+n+1} f(x)\,dx = \int_b^{b+1} f(y+n)\,dy = \int_b^{b+1} f(y)\,dy$$

were the last line follows from "1-periodicity" of f. If $b = 0$ we're done. Otherwise write

$$\begin{aligned}\int_b^{b+1} f(y)\,dy &= \int_b^1 + \int_1^{b+1} f(y)\,dy \\ &= \int_b^1 f(y)\,dy + \int_0^b f(z-1)\,dz \\ &= \int_b^1 f(y)\,dy + \int_0^b f(z)\,dz\end{aligned}$$

where we've applied the change of variable $z = y+1$ to the second integral in the second line, and used the periodicity of f in the last line. Putting it all together:

$$\int_a^{a+1} f(x)\,dx = \int_b^1 f(x)\,dx + \int_0^b f(x)\,dx = \int_0^1 f(x)\,dx.$$

which proves the Lemma, and with it, the Proposition. □

Corollary 6.3. *If f and g belong to L^1, then for a.e. $x \in \mathbb{R}$:*

$$\int_{-\pi}^{\pi} f(x-t)g(t)\,dt = \int_{-\pi}^{\pi} g(x-t)f(t)\,dt$$

It's an easy matter to check that L^1, with "convolution multiplication" is a ring. Corollary 6.3 shows that this ring is *commutative*: $f * g = g * f$ for all pairs of functions $f, g \in L^1$.

7 *Arithmetic means*

Given a sequence $(s_n)_0^\infty$ of complex numbers we can form the accompanying sequence $(\sigma_N)_0^\infty$ of *arithmetic means*:

$$\sigma_N = \frac{1}{N+1}\sum_{n=0}^{N} s_n \qquad (N = 0, 1, 2, \dots).$$

Abel's Theorem. *If the original sequence (s_n) converges to a complex number s, then so does its sequence of arithmetic means.*

It may happen that the sequence of arithmetic means converges even if the original sequence does *not*. For example, let $s_n = (-1)^n$. Then $\sigma_N = 0$ if N is odd, and $= \frac{1}{N+1}$ otherwise. Thus $\sigma_N \to 0$. This example suggests that arithmetic means tend to make oscillatory sequences more likely to converge. We'll now see that this is precisely what happens in the case of Fourier-series partial sums.

Definition 7.1. *For $f \in L^1$, $x \in \mathbb{R}$, and N a non-negative integer, let*

$$(\sigma_N f)(x) = \frac{1}{N+1}\sum_{n=0}^{N} (s_n f)(x)\,. \tag{14}$$

Thus $\sigma_N f$ is the N-th arithmetic mean of the sequence $(s_N f)_0^\infty$ of (symmetric) partial sums of the Fourier series of f. It is called the N-th *Fejér mean* of f. The goal of this section is to prove that if $f \in L^1$, then its sequence of Fejér means converges to f in the norm of L^1. More precisely:

Theorem 7.2 (Fejér's Theorem)**.** $\lim_{N\to\infty} \|\sigma_N f - f\|_1 = 0$ *for each* $f \in L^1$.

cf. Math. Ann. 58 (1904), pp. 51-69, where Fejér proved that if f is *continuous* and 2π-periodic on $\mathbb{R}$ then $\sigma_N f \to f$ *uniformly* on $\mathbb{R}$.

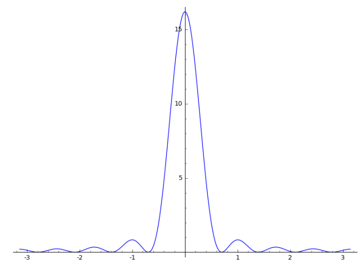
The proof requires some preliminaries, the first of which is a closed-form expression for Fejér means.

Proposition 7.3. *If $f \in L^1$, $x \in \mathbb{R}$, and $N = 0, 1, 2, \ldots$, then*

$$(\sigma_n f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) K_N(x-t)\, dt, \tag{15}$$

where K_N is the Fejèr kernel*, defined by*

$$K_N(x) = \frac{1}{N+1} \left(\frac{\sin\left((2N+1)\frac{x}{2}\right)}{\sin\left(\frac{x}{2}\right)} \right)^2 \tag{16}$$



The Fejér kernel K_4

Proof. By linearity of integral, Eqn. (15) holds with K_N the N-th arithmetic mean of the sequence of Dirichlet kernels:

$$K_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} D_n(x) = \frac{1}{N+1} \sum_{n=0}^{N} \frac{\sin\left((2n+1)\frac{x}{2}\right)}{\sin(\frac{x}{2})}$$

Thus

$$(N+1)\sin(\frac{x}{2}) K_N(x) = \sum_{n=0}^{N} \operatorname{Im} e^{i(n+\frac{1}{2})x} = \operatorname{Im} \sum_{n=0}^{N} e^{i(n+\frac{1}{2})x}. \tag{17}$$

Now

$$\begin{aligned}
\sum_{n=0}^{N} e^{i(n+\frac{1}{2})x} &= e^{i\frac{x}{2}} \sum_{n=0}^{N} e^{inx} = e^{i\frac{x}{2}} \cdot \frac{1-e^{i(N+1)x}}{1-e^{ix}} \\
&= e^{i\frac{x}{2}} \cdot \frac{e^{i\frac{N+1}{2}x}}{e^{i\frac{x}{2}}} \cdot \frac{e^{-i\frac{N+1}{2}x} - e^{i\frac{N+1}{2}x}}{e^{-i\frac{x}{2}} - e^{i\frac{x}{2}}} \\
&= e^{i\frac{N+1}{2}x} \cdot \frac{\sin\frac{N+1}{2}x}{\sin\frac{x}{2}}
\end{aligned}$$

Substitution of this last result into (17) yields:

$$(N+1)\sin(\frac{x}{2})\, K_N(x) = \left(\sin\tfrac{N+1}{2}x\right) \cdot \frac{\sin\frac{N+1}{2}x}{\sin\frac{x}{2}}$$

from which follows the promised formula for $K_N(x)$. □

Lemma 7.4. *The Fejér kernel sequence has the following properties:*

(a) $K_N(x) \geq 0$ *for* $N = 0, 1, 2, \ldots$ *and* $x \in \mathbb{R}$.

(b) $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x)\, dx = 1$ *for* $N = 0, 1, 2, \ldots$.

(c) $\lim_{N\to\infty} \max_{\delta \leq |x| \leq \pi} K_N(x) = 0$ *whenever* $0 < \delta \leq \pi$.

Proof. Property (a) is obvious from (16), while (b) follows immediately from the fact that $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)\, dx = 1$ for every N (as you can see by integrating the definition (10) of D_N, and using equation (1).

As for (c), we have from (16):

$$\max_{\delta \leq |x| \leq \pi} K_N(x) \leq \frac{1}{N+1} \cdot \frac{1}{\sin^2(\delta/2)} \to 0 \text{ as } N \to \infty. \qquad \square$$

Next, recall the notation f_t for the translate of the function f by $t \in \mathbb{R}$, as defined by Eqn. (13).

Lemma 7.5. $\lim_{t\to 0} \|f_t - f\|_1 = 0$ *for each* $f \in L^1$.

Proof. The result is easy if f is 2π-periodic and continuous (hence uniformly continuous) on the real line. From measure theory we know that in the space L^1 such functions form a dense subset.[6] Thus, given $\varepsilon > 0$ we can choose $g \in L^1$ continuous, such that $\|f - g\|_1 < \varepsilon/3$, and since the theorem is true for continuous functions, we can choose $0 < \delta < \pi$ so that

$$|t| < \delta \implies \|g - g_t\|_1 < \varepsilon/3\,.$$

Consequence: whenever $|t| < \delta$ we have

$$\|f - f_t\|_1 \le \underbrace{\|f - g\|_1}_{<\varepsilon/3} + \underbrace{\|g - g_t\|_1}_{<\varepsilon/3} + \underbrace{\|g_t - f_t\|_1}_{=\|g-f\|_1<\varepsilon/3} < \varepsilon\,,$$

where the identification of the last summand's norm comes from Proposition 6.1 and the fact that $f_t - g_t = (f - g)_t$. □

[6] See, [5, Theorem 11.28, pp. 326–7], where this is proved for the space L^2. The argument for L^1 is the same.

This is a classic "$\varepsilon/3$–argument," wherein one first proves the result for a dense subset, then uses some kind of uniform estimate to transfer this partial result to the whole space.

Proof of Theorem 7.2. Fix $f \in L^1$ and (for the moment) a non-negative integer N. Without loss of generality we can assume that $\|f\|_1 = 1$. By Corollary 6.3 we may interchange the roles of f and K_N in our convolution representation of the arithmetic means of the Fourier series of f, thus obtaining

$$(\sigma_N f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)K_N(t)\,dt \qquad (x \in \mathbb{R}).$$

By Lemma 7.4(a):

$$(\sigma_N f)(x) - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \big[f(x-t) - f(x)\big]K_N(t)\,dt \qquad (x \in \mathbb{R}).$$

Upon taking absolute values and using the non-negativity of the Fejér kernel, we obtain

$$|(\sigma_N f)(x) - f(x)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x-t) - f(x)\big|K_N(t)\,dt \qquad (x \in \mathbb{R}).$$

Integrate both sides of the above inequality on x, and use Fubini's Theorem to obtain

$$\|\sigma_N f - f\|_1 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} \|f_t - f\|_1 K_N(t)\,dt\,. \tag{18}$$

Let $\varepsilon > 0$ be given, and use Lemma 7.5 to choose $\delta \in (0,\pi)$ so that

$$|t| < \delta \implies \|f_t - f\|_1 < \varepsilon/2\,. \tag{19}$$

Now break the right-hand side of (18) into two pieces: I_δ where the integral extends over the interval $[-\delta, \delta]$, and J_δ where it extends over over $\{t \colon \delta \le |t| \le \pi\}$. We have

$$I_\delta \;\le\; \frac{\varepsilon}{2} \cdot \underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(t)\,dt}_{=1 \text{ by Lemma 7.4(b)}} \;=\; \frac{\varepsilon}{2},$$

and

$$\begin{aligned} J_\delta &= \frac{1}{2\pi}\int_{\delta\le|t|\le\pi} \|f_t - f\|_1\, K_N(t)\,dt \\ &\le \frac{1}{2\pi}\int_{\delta\le|t|\le\pi} (\|f_t\|_1 + \|f\|_1)\, K_N(t)\,dt\,. \end{aligned}$$

We know from the change-of-variable formula and the 2π-periodicity of f that $\|f_t\|_1 = \|f\|_1$, so the term in parentheses under the integral in the last line is just the constant $2\|f\|_1$. Thus

$$J_\delta \le 2\underbrace{\|f\|_1}_{=1} \cdot \frac{1}{2\pi}\int_{\delta\le|t|\le\pi} K_N(t)\,dt \le \frac{1}{\pi}\max_{\delta\le|t|\le\pi} K_N(t)$$

Use Lemma 7.4(c) to choose N_ε so that

$$N > N_\varepsilon \implies \max_{\delta\le|t|\le\pi} K_N(t) < \frac{\pi}{2}\,\varepsilon \implies J_\delta < \frac{\varepsilon}{2}\,.$$

Thus $N > N_\varepsilon$ implies

$$\|\sigma_N f - f\|_1 \;<\; I_\delta + J_\delta \;<\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \;=\varepsilon,$$

which completes the proof of Fejér's Theorem. □

8 *Consequences of Fejér's Theorem*

Corollary 8.1 (The Uniqueness Theorem for Fourier transforms)**.** *If f and g belong to L^1 and $\hat{f} \equiv \hat{g}$, then $f = g$ a.e. on $\mathbb{R}$.*

Proof. Since $\widehat{f-g} = \hat{f} - \hat{g}$ it's enough replace f by $f-g$ and g by the zero-function, i.e., we may without loss of generality assume that $f \in L^1$ with $\hat{f} \equiv 0$, and prove that $f = 0$ a.e..

So suppose that $\hat{f}(n) = 0$ for each $n \in \mathbb{Z}$. This implies for each non-negative integer N that the Fourier series partial sum $s_N f$ is the zero-function on $\mathbb{R}$, hence the same is true for the Fejér means $\sigma_N f$. By Theorem 7.2 this latter sequence (of zero-functions) converges in L^1 to f, hence $f = 0$ a.e. on $\mathbb{R}$. □

The Uniqueness Theorem asserts that the *Fourier transform* $\mathscr{F} \colon f \to \hat{f}$, which we observed takes L^1 linearly into $\ell^\infty(\mathbb{Z})$ is one-to-one, thus answering the first part of Question Q2 posed on page 3.

Another consequence of Fejér's Theorem: *Every L^1-function is the limit, in the norm of L^1, of a sequence of trigonometric polynomials*—namely, the arithmetic means of the symmetric partial sums of its Fourier series. Qualitatively:

Corollary 8.2. *The trigonometric polynomials form a dense subspace of L^1.*

We've shown earlier that the Fourier transform takes L^1 into the space $\ell^\infty(\mathbb{Z})$ (those two-sided complex sequences that are *bounded*), and commented that it actually maps L^1 into c_0 (those two-sided bounded complex sequences that converge to zero as $|n| \to \infty$). Thanks to Fejer's Theorem, we can now prove this.

Corollary 8.3 (The Riemann-Lebesgue Lemma). $\lim_{|n|\to\infty} \hat{f}(n) = 0$ *for each* $f \in L^1$.

Proof. Fix $f \in L^1$. The result clearly holds if f is a trigonometric polynomial, in which case $\hat{f}(n) = 0$ for all sufficiently large $|n|$. Let $\varepsilon > 0$ be given. By Corollary 8.2 there is a trigonometric polynomial p with $\|f - p\|_1 < \varepsilon$. We may therefore choose a positive integer N so that $\hat{p}(n) = 0$ for $|n| > N$. Thus $|n| > N$ implies:

$$|\hat{f}(n)| = |\hat{f}(n) - \hat{p}(n)| = |\widehat{(f-p)}(n)| \le \|f - p\|_1 < \varepsilon.$$

where the second equality results from the linearity of the Fourier transform, and the "contractivity" (7) of the Fourier transform.

Conclusion: $\hat{f}(n) \to 0$ as $|n| \to \infty$, as we wished to show. □

9 *Remarks on the proof of Fejér's Theorem*

9.1 *Convergence and continuity*

In the proof of Theorem 7.2, the "$\frac{\varepsilon}{2}$-argument" that followed inequality (18) used only that the function F defined by

$$F(t) = \|f_t - f\|_1 \qquad (t \in \mathbb{R})$$

is continuous at the origin, and vanishes there. Thus it proves the following corollary (of the proof of Theorem 7.2):

Corollary 9.1. *If* $f \in L^1$ *is continuous at a point* $x_0 \in \mathbb{R}$, *then* $(\sigma_N f)(x_0) \to f(x_0)$ *as* $N \to \infty$.

Proof. Apply the argument following (18) with $\|f_t - f\|_1$ replaced by $|f(t - x_0) - f(x_0)|$. □

If, in addition, we assume that f is continuous *at each point of* $\mathbb{R}$ (hence, by 2π-periodicity, *uniformly continuous* there), the same argument yields Fejér's original result: $\sigma_N f \to f$ *uniformly* on $\mathbb{R}$. In the same way that our L^1 Fejér theorem yields Corollary 8.2, this "continuous" version yields a famous result due to Weierstrass:

Corollary 9.2 (The Weierstrass (trigonometric) Approximation Theorem)**.** *If f is continuous and 2π-periodic on the real line, then there is a sequence $(p_n)_0^\infty$ of trigonometric polynomials that converges uniformly on the line to f.*

9.2 *Approximate identities*

The proof of Fejér's Theorem works for any sequence of functions $(f * k_N)_0^\infty$, where $f \in L^1$ and $(k_N)_0^\infty$ is a sequence of L^1 functions satisfying the properties (a), (b), and (c) listed for the Fejér-kernel sequence in the statement of Lemma 7.4. More precisely we have:

Theorem 9.3. *Suppose $(k_N)_0^\infty$ is a sequence of functions in L^1 satisfying conditions (a), (b), and (c) of Lemma 7.4. Then $\lim_{N\to\infty} \|k_N * f - f\|_1 \to 0$ for every $f \in L^1$.*

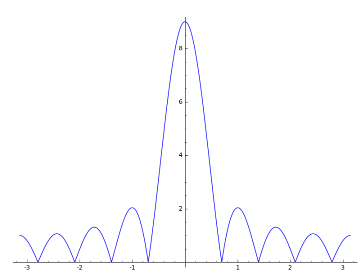
For this reason, such a sequence (k_N) is called an *approximate identity* for L^1. It's an interesting exercise to show that, with convolution as multiplication, L^1 is a commutative ring *that has no identity function*, i.e., no function e such that $f * e = f$ for each $f \in L^1$.

In the proof of Theorem 7.2 we did not need the full strength of property (c) of the Fejér-kernel sequence. It would have been enough to require :

(c′) $\lim_{N\to\infty} \int_{\delta \le t \le \pi} K_N(t)\, dt = 0$ for each $\delta \in (0, \pi)$.

Consequently, for Theorem 9.3 one can extend the definition of "approximate identity" sequence $(k_N)_0^\infty$ require only conditions (a) and (b) of Lemma 7.4, and (c′) above.

9.3 *What's wrong with the Dirichlet kernel?*



The 'absolute" Dirichlet kernel D_4

To get a feeling for why the partial sum sequence of a Fourier series behaves so much worse than the Fejér-kernel sequence, one need only look at the graph of the absolute values of the functions in the Dirichlet-kernel sequence (see the figure at the right for the graph of $|D_4|$). In general the graph of $|D_N|$ contains of a large central "hump" and $(N-1)$ smaller ones, on each side of the origin, the k-th of which (call it H_k) on the positive side has base of length $b_k = \pi/(2N+1)$ and height

$$h_k = \left|\sin\left(\frac{(2k+1)\pi}{2N+1}\right)\right|^{-1} \ge \frac{2N+1}{(2k+1)\pi}.$$

Here we use the inequality $\sin\theta \le \theta$ for $0 \le \theta \le \pi/2$.

Since H_k is a convex curve, the region it bounds contains the isosceles triangle Δ_k that shares its base with H_k and has height h_k. Thus the area of H_k is $>$ the area, $\frac{1}{2}b_k h_k \geq c/(2k+1)$, of Δ_k, hence

$$\|D_N\|_1 > 2\sum_{k=1}^{N-1} \text{Area of } \Delta_k \geq \sum_{k=1}^{N-1} \frac{1}{2k+1} \geq \text{const.} \log N\,,$$

where "const." is a positive constant that does not depend on N. *Conclusion*: $\|D_N\|_1 \to \infty$ as $N \to \infty$.

The contrast between this L^1-unboundedness for the Dirichlet-kernel sequence and the corresponding boundedness of the Fejér-kernel sequence (property (b) in the statement of Lemma 7.4) turns out to be the root of all difficulties associated with the partial-sum sequences of Fourier series.

In fact, the argument here shows that, in sharp contrast with property (c') above in the "improved" definition of "approximate identity,"

$$\lim_{N\to\infty} \int_{\delta\leq|z|\leq\pi} |D_N(x)|\, dx = \infty$$

for each $\delta \in (0, \pi)$.

References

1. Radomir S. Stanković, Jaakko T. Astola, and Mark G. Karpovsky, *Remarks on the history of abstract harmonic analysis*, A beautiful set of lecture slides, freely downloadable from:

   `http://www.cs.tut.fi/~jta/computing-history-material/fourierhistory.pdf`.

2. Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Mathematica 116 (1966) 135–157.

3. Leopold Fejér, *Untersuchungen über Fouriersche Reihen*, Mathematische Annalen 58 (1904) 51-69.

4. Andrei N. Kolmogorov, *Une série de Fourier-Lebesgue divergente presque partout*, Fundamenta Math. 4 (1923) 324–328.

5. Walter Rudin, *Principles of Mathematical Analysis*, 3rd ed. McGraw-Hill 1976

6. Walter Rudin, *Real and Complex Analysis*, 3rd ed., McGraw-Hill 1987.