Appendix B

More on the Convergence of Fourier Series

We return to the big, unanswered question: Is the Fourier series

$$\sum_{n=0}^{\infty}\left[a_n\cos(nx)+b_n\sin(nx)\right],$$

with the coefficients calculated from a function f by the proper formulas, actually equal to $f(x)$? In what sense, and under what circumstances?

We will discuss four different types of convergence behavior:

1. pointwise convergence
2. uniform convergence
3. summability
4. mean convergence

In each case we have to define what the term means and learn conditions that are adequate to assure that it applies to the function in question. Along the way we'll seek out examples of functions to which it *doesn't* apply, the better to appreciate the property when we have it and to understand why such conditions are needed to insure it.

The philosophy of this course is that you should see proofs of the major Fourier convergence theorems — primarily for the "character-building" effects of the experience. Students who are not math majors seldom take upper-level math courses that, for the first time, pay serious attention to proofs. But competent physicists (etc.) need to develop some degree of "literacy" about what is involved in rigorous mathematics and why; this topic is perhaps the best and last opportunity. You are not expected to memorize the proofs and regurgitate them on tests. You are expected to pay attention to the lectures and read the books (especially Tolstov) to fill in and reinforce the lectures. There will be a big homework assignment in which you will be led to construct a proof of the summability theorem by making minor changes in the proof of the pointwise convergence theorem that I'm about to present.

Once we have built up an understanding of the Fourier theory, in the second half of the course we'll take for granted the similar theorems that apply in more complicated situations, without looking at the proofs at all.

Of course, much more important than the details of proofs is an understanding of what the theorems say.

"Pointwise convergence" means that we are applying the standard (2nd-semester calculus) definition of convergence of a sequence of *numbers* to the values of our functions and series at each point x in the domain.

Definitions: The Mth *partial sum* of the Fourier series is

$$S_M(x) \equiv \sum_{n=0}^{M} \bigl[a_n \cos(nx) + b_n \sin(nx)\bigr].$$

The series *converges pointwise* to f (a periodic function with period 2π) if

$$f(x) = \lim_{M\to\infty} S_M(x) \quad \text{for every } x.$$

That is, for any x and for any $\epsilon > 0$ there is an N such that

$$|f(x) - S_M(x)| < \epsilon \quad \text{for all } M > N.$$

The strategy of the proof will slowly emerge as I gather together some technical machinery; please be patient.

Lemma 1: $1 + 2\sum_{n=1}^{M} \cos(nx) = \dfrac{\sin\bigl[(M+\frac{1}{2})x\bigr]}{\sin\frac{x}{2}}$
for positive integer M.

Proof: Use $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ repeatedly:

$$\sin\bigl(M \pm \tfrac{1}{2}\bigr)x = \sin(Mx)\cos\frac{x}{2} \pm \cos(Mx)\sin\frac{x}{2}.$$

Subtract:

$$\sin\bigl(M + \tfrac{1}{2}\bigr)x = 2\cos(Mx)\sin\frac{x}{2} + \sin\bigl(M - \tfrac{1}{2}\bigr)x.$$

Treat the last term in the same way [note: $M - \frac{1}{2} = (M-1) + \frac{1}{2}$], getting

$$\sin\bigl(M + \tfrac{1}{2}\bigr)x = 2\cos(Mx)\sin\frac{x}{2} + 2\cos[(M-1)x]\sin\frac{x}{2} + \sin\bigl(M - \tfrac{3}{2}\bigr)x.$$

Continue in this way until the frequency factor in the last term hits bottom:

$$\sin\bigl(M + \tfrac{1}{2}\bigr)x = 2\sum_{n=1}^{M} \cos(nx)\sin\frac{x}{2} + \sin\frac{x}{2}.$$

Divide by $\sin\frac{x}{2}$ to get the desired identity.

Remark: The denominator vanishes if x is a multiple of 2π, but so does the numerator. We can define the function at such a point as the limit of the fraction — which is $1+2M$ by L'Hôpital's rule.

Lemma 2: $\displaystyle\int_0^\pi \frac{\sin\left(M+\frac{1}{2}\right)x}{\sin\frac{x}{2}}\,dx = \pi = \int_{-\pi}^0 \frac{\sin\left(M+\frac{1}{2}\right)x}{\sin\frac{x}{2}}\,dx.$
(In particular, these integrals are independent of M.)

Proof: Integrate Lemma 1:

$$\begin{aligned}\int_{-\pi}^\pi \frac{\sin\left(M+\frac{1}{2}\right)x}{\sin\frac{x}{2}}\,dx &= \int_{-\pi}^\pi \left(1+2\sum_{n=1}^M \cos(nx)\right)dx \\ &= 2\pi\end{aligned}$$

by orthogonality. Since the integrand is even, half the integral comes from $0<x<\pi$ and half from $-\pi<x<0$.

Lemma 3: $\displaystyle S_M(x) = \frac{1}{2\pi}\int_{-\pi}^\pi f(x+u)\,\frac{\sin\left[\left(M+\frac{1}{2}\right)u\right]}{\sin\frac{u}{2}}\,du.$
(Here it's understood that f is a function with period 2π and the coefficients in the Fourier sum S_M are computed from f by the standard Fourier formulas, $a_0 = \frac{1}{2\pi}\int_{-\pi}^\pi f(t)\,dt$, etc.)

Proof: Substitute the coefficient formulas into the series, taking care to use a different letter for the integration variable:

$$\begin{aligned}S_M(x) &= \frac{1}{2\pi}\int_{-\pi}^\pi f(t)\,dt \\ &\qquad + \frac{1}{\pi}\sum_{n=1}^M \left[\int_{-\pi}^\pi f(t)\cos nt\,dt\,\cos nx + \int_{-\pi}^\pi f(t)\sin nt\,dt\,\sin nx\right] \\ &= \frac{1}{\pi}\int_{-\pi}^\pi f(t)\left[\frac{1}{2}+\sum_{n=1}^M(\cos nt\cos nx+\sin nt\sin nx)\right]dt \\ &= \frac{1}{\pi}\int_{-\pi}^\pi f(t)\left[\frac{1}{2}+\sum_{n=1}^M \cos[n(t-x)]\right]dt\end{aligned}$$

Let $t = x+u$:

$$S_M(x) = \frac{1}{2\pi}\int_{-\pi}^\pi f(x+u)\left(1+2\sum_{n=1}^M \cos nu\right)du,$$

which is what we want, according to Lemma 1.

Now it should be clear what we've been up to. We have rewritten the partial sum as an integral involving a fairly simple function, whose behavior as $M \to \infty$ we can easily study. The *Dirichlet kernel*,

$$D_M(u) = \frac{\sin\left[\left(M + \frac{1}{2}\right)u\right]}{2\pi \sin \frac{u}{2}},$$

has a central peak that becomes tall and narrow when M becomes large. Therefore, the integral in Lemma 3 samples the values of $f(x+u)$ only for $u \approx 0$. So the idea is that when $M \to \infty$ the kernel acts as a *delta function*, so that the integral approaches $f(x)$.

More precisely, we will prove that if f satisfies certain conditions, then

$$\begin{aligned}\lim_{M\to\infty} \int_0^\pi f(x+u) D_M(u)\, du &= \frac{1}{2} \lim_{u \downarrow 0} f(x+u) \\ &= \frac{1}{2} \lim_{t \downarrow x} f(t) \equiv \frac{1}{2} f(x+0)\end{aligned}$$

and

$$\begin{aligned}\lim_{M\to\infty} \int_{-\pi}^0 f(x+u) D_M(u)\, du &= \frac{1}{2} \lim_{u \uparrow 0} f(x+u) \\ &\equiv \frac{1}{2} f(x-0).\end{aligned}$$

Then Lemma 3 says that

$$\boxed{\lim_{M\to\infty} S_M(x) = \tfrac{1}{2}\left[f(x-0) + f(x+0)\right].}$$

The significance of this equation depends on how f itself behaves at x. Consider these cases:

1. *Continuity:* $f(x) = f(x-0) = f(x+0)$. Then $\lim_{M\to\infty} S_M(x) = f(x)$, which is what we've been trying to prove. The Fourier series converges to the value of the function at x.

2. *Removable discontinuity:* $f(x-0) = f(x+0)$, but either $f(x)$ is undefined or $f(x) \neq f(x+0)$. (This might happen when f is originally defined only on the open interval $(0,\pi)$ and then we pass to its odd periodic extension.) In this case the series converges to a value that fills in a hole in the graph of f in a natural way. Often one then defines (or redefines) $f(x)$ to be equal to $f(x+0) = f(x-0)$. (Changing a function at one point, or a finite number of points, can't change the values of integrals involving the function; this includes the Fourier coefficients a_n and b_n.)

3. *Jump discontinuity:* $f(x-0) \neq f(x+0)$, but both limits do exist. Then the series converges to the average of the two limits, $\frac{1}{2}[f(x-0)+f(x+0)]$. Again, without changing the series one can redefine $f(x)$ to equal this compromise value. This new f is *not* continuous, but the series does converge pointwise to $f(x)$ everywhere.

4. *Worse discontinuities:* On or the other of the one-sided limits does not exist. Examples:

   a) $\dfrac{1}{x}$ (at 0);

   b) $\sin\dfrac{1}{x}$ (at 0);

   c) $f(x) = \begin{cases} 1 & \text{if } x \text{ is rational,} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$ (anywhere).

   We will not consider such cases; the hypotheses on f in the theorem will exclude them.

We are now ready to frame some hypotheses on f that will enable us to finish the proof of the basic formula in the box above.

Definition: A function defined on the circle is *piecewise continuous* (or *sectionally continuous*) if it has (at most) a finite number of jumps and removable discontinuities, and no worse discontinuities. A periodic function is piecewise continuous if it has at most finitely many such discontinuities in each period. A function defined on the interval $(-\pi, \pi)$ is piecewise continuous if it has at most finitely many such discontinuities *and* the endpoint limits, $f(\pi-0)$ and $f(-\pi+0)$, exist. (If the two endpoint limits are not equal, this counts as a jump discontinuity.)

Definition: f is *piecewise smooth* if not only f but also f' is piecewise continuous. (Thus f may have finitely many jumps and finitely many corners (jumps of f'), but no vertical tangents (worse discontinuities of f').)

For example, $|x|$ is piecewise smooth, but $\sqrt[3]{x}$ and $\sqrt{|x|}$ are not.

Riemann–Lebesgue Theorem: If $\phi(x)$ is piecewise continuous on $a \leq x \leq b$, then

$$\int_a^b \phi(x)\sin(kx)\,dx \to 0 \quad \text{and} \quad \int_a^b \phi(x)\cos(kx)\,dx \to 0 \quad \text{as } k \to \infty,$$

where k is a real variable.

I shall postpone the proof of this theorem until after we have used it in the proof of the pointwise convergence theorem, to which we now return. The next lemma looks rather obvious, but it needs to be proved.

Lemma 4: If f is piecewise smooth and x is a point of discontinuity of f or f', then the *right derivative* of f at x, defined as

$$\lim_{h\downarrow 0} \frac{f(x+h)-f(x+0)}{h},$$

exists and equals $f'(x+0)$. [In $f'(x+0)$ the limit defining the derivative is taken at a point to the right of x, and then that point is made to approach x. In the right derivative, the difference quotient is set up right at x from the start.] Similarly, the *left derivative* equals $f'(x-0)$. (Of course, if x is a point of continuity of f and f', then the right and left derivatives are just equal to $f'(x)$ and there is nothing to make a fuss over.)

Proof: By the mean value theorem,

$$\frac{f(x+h)-f(x+0)}{h} = f'(x+k) \quad \text{for some } k \text{ between } 0 \text{ and } h.$$

Since f' is piecewise continuous, $\lim_{k\downarrow 0} f'(x+k) = f'(x+0)$, QED.

Now we can state and (almost) prove the key lemma from which, as I recently explained, the pointwise convergence theorem follows.

Lemma 5: If f is piecewise smooth, then

$$\lim_{M\to\infty} \int_0^\pi f(x+u) D_M(u)\,du = \frac{1}{2} f(x+0),$$

$$\lim_{M\to\infty} \int_{-\pi}^0 f(x+u) D_M(u)\,du = \frac{1}{2} f(x-0).$$

Proof: Obviously the two halves of the lemma will have almost identical proofs, so we concentrate on the first half. By Lemma 2,

$$\begin{aligned}\frac{1}{2} f(x+0) &= \frac{1}{2\pi} \int_0^\pi \frac{\sin(M+\frac{1}{2})u}{\sin\frac{u}{2}}\,du\, f(x+0) \\ &= \int_0^\pi D_M(u)\,du\, f(x+0).\end{aligned}$$

So the lemma is equivalent to

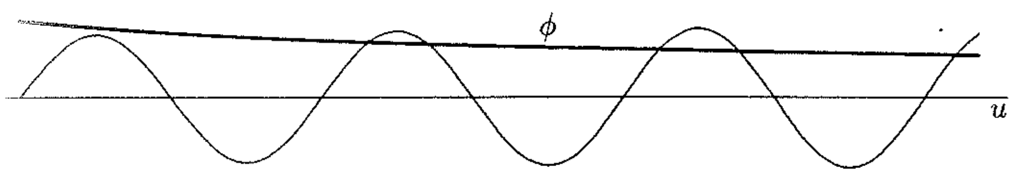
$$\lim_{M\to\infty} \int_0^\pi [f(x+u)-f(x+0)] \frac{\sin(M+\frac{1}{2})u}{2\pi\sin\frac{u}{2}}\,du = 0.$$

This integral has the form

$$\int_0^\pi \phi(u) \sin\left(M + \tfrac{1}{2}\right)u \, du,$$

with

$$\phi(u) \equiv \frac{1}{\pi} \frac{\frac{u}{2}}{\sin \frac{u}{2}} \frac{f(x+u) - f(x+0)}{u}.$$



As M becomes large, $\sin\left[\left(M + \frac{1}{2}\right)u\right] \equiv \sin t$ oscillates increasingly rapidly. Eventually $\phi(u) \equiv \phi\left(t/(M + \frac{1}{2})\right)$ will vary so slowly over one period of the sine that the positive and negative parts of the integral will almost cancel. So we expect that the limit will be 0, as required. Indeed, the Riemann–Lebesgue theorem guarantees this, if ϕ is piecewise continuous. Well, ϕ is obviously piecewise continuous everywhere except possibly at $u = 0$. There, we have to examine each factor carefully. First,

$$\frac{\frac{u}{2}}{\sin \frac{u}{2}}$$

approaches 1 as $u \to 0$ (and the other zeros of $\sin \frac{u}{2}$ are outside the interval of interest). Second,

$$\frac{f(x+u) - f(x+0)}{u} \to f'(x+0) \quad \text{as } u \downarrow 0,$$

since f is piecewise smooth (Lemma 4). Therefore, $\lim_{u \downarrow 0} \phi(u)$ exists, and hence ϕ is piecewise continuous.

So, except for taking the Riemann–Lebesgue theorem on faith, we have finally finished the proof of ...

Pointwise convergence theorem for Fourier series: If f has period 2π and is piecewise smooth, then at each point x its Fourier series converges to

$$\frac{1}{2}[f(x-0) + f(x+0)].$$

In particular, it converges to $f(x)$ if f is continuous at x.

Corollaries and other remarks:

1. If f is given on $(-\pi, \pi)$, then at the endpoints the series converges to

$$\frac{1}{2}[f(\pi - 0) + f(-\pi + 0)].$$

2. Obviously the theorem holds for functions of arbitrary period $2L$, or for Fourier series on an arbitrary interval of length $2L$. (Just change n to $\frac{n\pi}{L}$ in all formulas.)

3. Sine series and cosine series on intervals of length L converge, for piecewise smooth f. (Consider the odd or even extension.)

4. The theorem says nothing about what happens if f is *not* piecewise smooth. The series *may* still converge in such a case. There are other theorems guaranteeing convergence for various slightly rougher kinds of functions. There are also examples of slightly rougher functions for which the series does not converge.

Proof of the Riemann–Lebesgue theorem

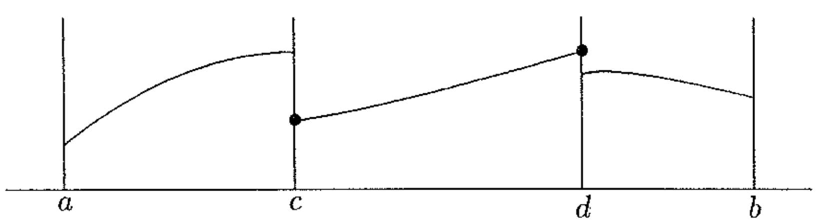
This theorem was stated earlier (after definition of "piecewise continuous"), and the intuitive idea of the proof was indicated in passing in the proof of Lemma 5.

We want to prove that

$$\int_a^b \phi(x)\,\sin(kx)\,dx \to 0 \quad \text{as } k \to \infty$$

(and the corresponding statement for the cosine, whose proof is the same). The hypothesis is that ϕ is piecewise continuous on $[a, b]$.

The integral is the sum of finitely many (maybe just one) integrals over smaller intervals where ϕ is continuous, so we need only to prove that the integral over *each such segment* approaches 0.



Without changing the integral from c to d, we may replace $\phi(c)$ by $\phi(c+0)$ and $\phi(d)$ by $\phi(d-0)$ while treating that segment; so we can assume that ϕ is continuous on the *closed* interval $[c, d] \equiv I$. (Note: The possibility of doing this is built into our assumption that the function has "no worse singularities than jumps". It is not true for all functions.) Therefore, we can use a famous (but hard to prove) theorem:

Lemma 6: A continuous function ϕ on a closed interval I is *bounded* and *uniformly continuous* there. That is,

(a) There is a number M such that $|\phi(x)| \le M$ for all x in I.

(b) For every $\epsilon > 0$ there is a $\delta > 0$ such that for all x and y in I,

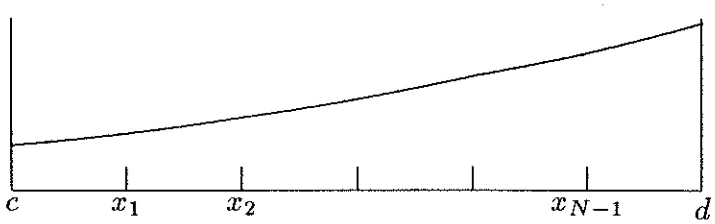
$$|\phi(x) - \phi(y)| < \epsilon \quad \text{whenever } |x - y| < \delta.$$

This is a theorem of general advanced calculus, not Fourier analysis, so we will not prove it here.

Now look at

$$\left| \int_c^d \phi(x) \sin(kx)\, dx \right|.$$

We must show that it is small if k is large.



Chop the interval into N small pieces, with endpoints $x_0 \equiv c, x_1, \ldots, x_{N-1}, x_N \equiv d$, as if we were forming a Riemann sum. Let ξ_i be any point between x_{i-1} and x_i. Then

$$\begin{aligned}
\left| \int_c^d \phi(x) \sin(kx)\, dx \right| &= \left| \sum_{i=1}^N \int_{x_{i-1}}^{x_i} \phi(x) \sin(kx)\, dx \right| \\
&\le \sum_i \left| \int_{x_{i-1}}^{x_i} \phi(x) \sin(kx)\, dx \right| \\
&= \sum_i \left| \int_{x_{i-1}}^{x_i} [\phi(x) - \phi(\xi_i)] \sin(kx)\, dx + \int_{x_{i-1}}^{x_i} \phi(\xi_i) \sin(kx)\, dx \right| \\
&\le \sum_i \left| \int_{x_{i-1}}^{x_i} [\phi(x) - \phi(\xi_i)] \sin(kx)\, dx \right| + \sum_i \left| \int_{x_{i-1}}^{x_i} \phi(\xi_i) \sin(kx)\, dx \right| \\
&\equiv A + B.
\end{aligned}$$

We will show that A and B can be made small simultaneously:

$$
\begin{aligned}
A &\le \sum_i \int_{x_{i-1}}^{x_i} |\phi(x) - \phi(\xi_i)|\,|\sin(kx)|\,dx \\
&\le \sum_i \int_{x_{i-1}}^{x_i} |\phi(x) - \phi(\xi_i)|\,dx \\
&\le \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} \frac{\epsilon}{2(d-c)} \quad \text{if the } x_i \text{ are close enough together.}
\end{aligned}
$$

(Lemma 6(b) says that given ϵ, we can choose the x_i close enough to make the last inequality true.) But

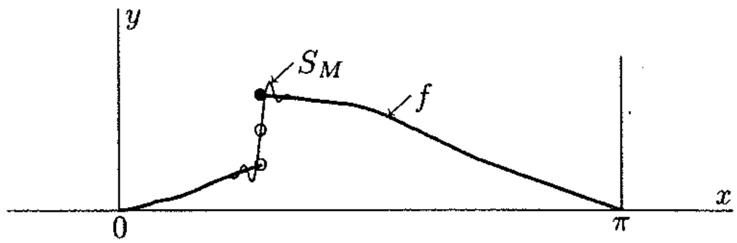
$$
\sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} \frac{\epsilon}{2(d-c)} = \frac{\epsilon}{2(d-c)} \int_c^d dx = \frac{\epsilon}{2}\,.
$$

This choice of the x_i may make N very large, but N does not depend on k. That is, the same N works for all k, so we can assume that N is fixed as we study the behavior of B at large k.

$$
\begin{aligned}
B &\le \sum_i |\phi(\xi_i)| \left| \frac{\cos(kx_{i-1}) - \cos(kx_i)}{k} \right| \\
&\le \frac{2}{k} \sum_{i=1}^{N} |\phi(\xi_i)| \\
&\le \frac{2MN}{k} \quad \text{(by Lemma 6(a)).}
\end{aligned}
$$

As $k \to \infty$ we eventually have $B < \epsilon/2$ and hence $A + B < \epsilon$, which is what we want to prove.

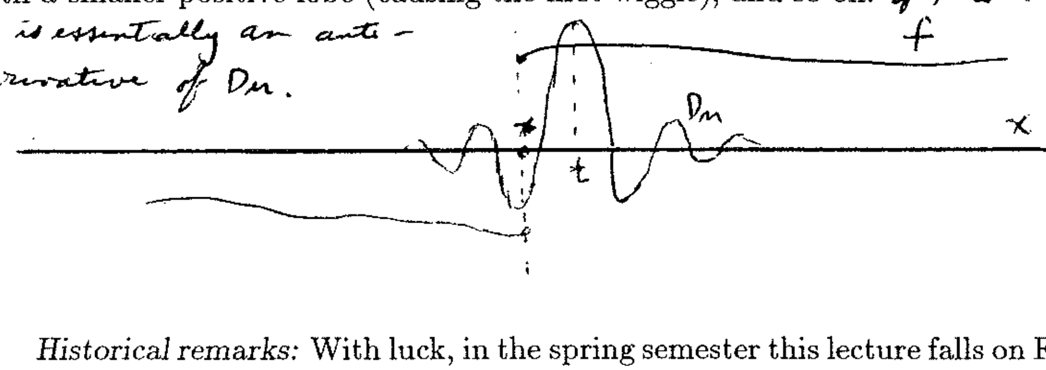
The pointwise convergence theorem says that, *if you wait long enough* (take enough terms in the sum), the Fourier series will converge even at and nearby a jump discontinuity. However, near a jump discontinuity any particular partial sum, S_M, is not as good an approximation to the function as it is elsewhere. It wiggles a lot, and actually develops "ears" or "overshoots" right next to the jump. This is called the *Gibbs phenomenon.* The height of the overshoot is typically about 10% of the size of the jump. The simplest example is provided by a square wave, such as an odd periodic function that is constant on $(0, \pi)$; see p. 66 of Powers.



Why does this happen? Essentially, the wiggles in the partial sum are just the wiggles in the Dirichlet kernel, D_M. Recall that

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

The central peak in the kernel "samples" f near x. As the kernel slides along the t axis just to the right of x, it samples the graph just to the left of x first with the central peak (causing the plunge of the graph of S_M down to the middle of the gap in the graph of f), then with a large negative lobe (causing the overshoot), then with a smaller positive lobe (causing the first wiggle), and so on. If f is a step function, S_M is essentially an anti-derivative of D_M.



Historical remarks: With luck, in the spring semester this lecture falls on February 11, the birthday of J. Willard Gibbs, the first great American theoretical and mathematical physicist. Gibbs's observation about Fourier series was a very minor incident in his career. It appears in a letter to *Nature* in 1899 (volume **59**, page 606). Gibbs had stepped into a controversy between the physicist Michelson and the mathematician Love over whether the graph of a square wave should include the vertical segments usually drawn in physics and engineering texts (which clearly can't be part of the graph of a function). These earlier letters, which make interesting reading, are in *Nature* **58**, 544–545; **58**, 569–570; **59**, 200–201 (1898). Gibbs's final conclusion was that the *limit of the graphs* of the partial sums (as opposed to the *graph of the limit function*) includes not only the vertical segments but also short extensions of them — the limits of the Gibbs overshoots. See also the historical article by H. S. Carslaw, *Bull. Amer. Math. Soc.* **31**, 420–424 (1925), which says that the phenomenon was actually discovered by someone named Wilbraham in 1848!

Uniform convergence

More important than the Gibbs phenomenon is *nonuniform convergence*, which always occurs near a jump. This is simply the problem we already noticed: Near a jump in f, any particular partial sum of the Fourier series is not really a very good approximation to f. Intuitively, S_M (which is continuous) has trouble becoming steep enough to approximate the jump. This would be true even if there were no Gibbs overshoot.

Definition: Let $\{S_M(x)\}$ be any *sequence of functions* defined on an interval $a \le x \le b$. [Usually in applications the functions S_M will be the partial sums of a *series of functions*:

$$S_M(x) \equiv \sum_{n=1}^{M} w_n(x)\,.]$$

$S_M(x)$ *converges uniformly* to a limit function $f(x)$ if

$$\max_{a\le x\le b}\bigl|f(x)-S_M(x)\bigr| \to 0 \quad \text{as } M\to\infty.$$

This is very close to what we usually mean by saying that S_M is guaranteed to be a *good approximation* to f if M is big. That is, in practice this is a property that you would very much like your approximations to have as often as possible!

To compare and contrast uniform convergence with pointwise convergence, note that uniform convergence can be reexpressed thus:

For every $\epsilon > 0$ there is an N such that for all x in the interval,

$$\bigl|f(x)-S_M(x)\bigr| < \epsilon \quad \text{if } M > N.$$

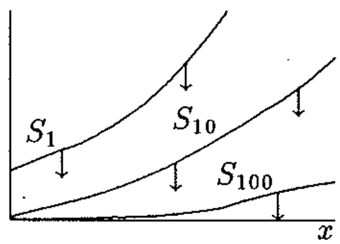
And remember that pointwise convergence means

For every x in the interval and for every $\epsilon > 0$ there is an N such that

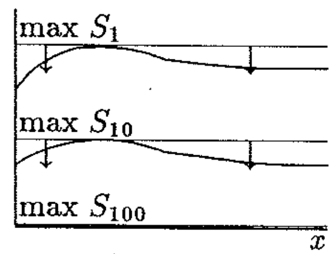
$$\left|f(x) - S_M(x)\right| < \epsilon \quad \text{if } M > N.$$

The crucial difference is that in the second case, N *is allowed to depend on* x, while in the first case, *the same* N *works for all* x (but depends on ϵ, of course). Therefore, if a sequence or series converges uniformly, then it also converges pointwise; but it may converge pointwise without converging uniformly.

[The same sort of distinction is involved in the definition of "uniform continuity" (see Lemma 6(b)). Conpare that with the definition of ordinary "continuity" in your calculus textbook.]



pointwise convergence to $f(x) = 0$

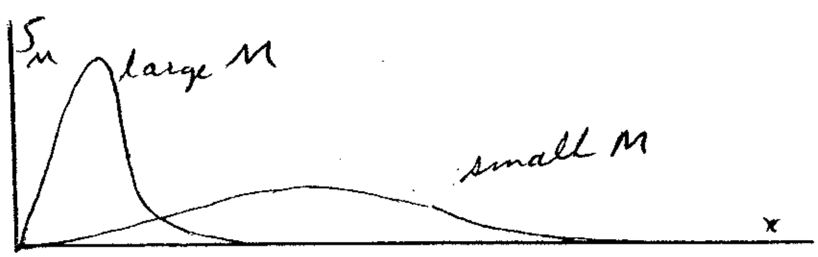


uniform convergence to 0

An example of a sequence that converges pointwise but *not* uniformly is

$$S_M(x) \equiv Mxe^{-Mx^2}.$$

Pointwise, $S_M(x) \to 0$ for all x, since the decaying exponential will eventually overcome x or any other power. But for a given M, consider $x = 1/\sqrt{M}$: Then $S_M(x) = \sqrt{M}\, e^{-1}$, which grows with M. Thus the maximum deviation of $S_M(x)$ from the limit function, $f(x) \equiv 0$, does *not* go to 0.



Why do we care? For one thing, uniform convergence is a hypothesis in many useful theorems. For example, ...

Theorem: A uniformly convergent series of continuous functions can be integrated (over a finite interval) term-by-term:

$$\int_a^b \sum_{n=1}^{\infty} w_n(x)\,dx = \sum_{n=1}^{\infty} \int_a^b w_n(x)\,dx.$$

Proof: Let $f(x) \equiv \sum_{n=1}^{\infty} w_n(x)\,dx$, and let $\{S_M\}$ be the partial sums, as usual. Then the theorem can be restated this way:

$$\int_a^b f(x)\,dx = \lim_{M\to\infty} \int_a^b S_M(x)\,dx.$$

But

$$\begin{aligned}\left|\int f\,dx - \int S_M\,dx\right| &\le \int |f - S_M|\,dx \\ &\le (b-a)\max(f - S_M) \\ &\to 0.\end{aligned}$$

To see that the uniformity assumption is needed in the theorem, observe that our example $S_M(x) = Mxe^{-Mx^2}$ violates the conclusion. We have

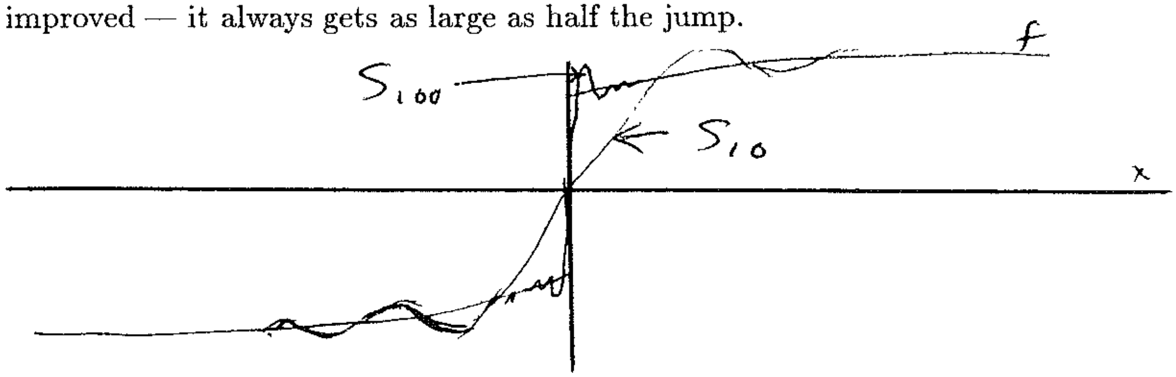
$$\int_0^1 f(x)\,dx = \int_0^1 0\,dx = 0,$$

but

$$\begin{aligned}\int_0^1 S_M(x)\,dx &= \int_0^1 Mxe^{-Mx^2}\,dx \\ &= \frac{1}{2}\int_0^M e^{-t}\,dt \\ &= \tfrac{1}{2} - \tfrac{1}{2}e^{-M} \\ &\to \tfrac{1}{2} \neq 0.\end{aligned}$$

(If you integrate to ∞ instead of 1, you get exactly $\frac{1}{2}$; the area under the curves is independent of M.)

Now let's get back to Fourier series. It is easy to see that if f is discontinuous, then uniform convergence of its Fourier series is impossible, because any continuous approximating function S_M "needs time" to cross the gap. There will always be points near the jump point where the approximation is bad. Taking M larger makes this bad interval narrower, but the magnitude of the error *inside* the interval is not improved — it always gets as large as half the jump.



In fact, as a general theorem, a discontinuous function can't be the *uniform* limit of a sequence or series of continuous functions:

Theorem: If a sequence of continuous functions S_M converges *uniformly* to f, then f is continuous.

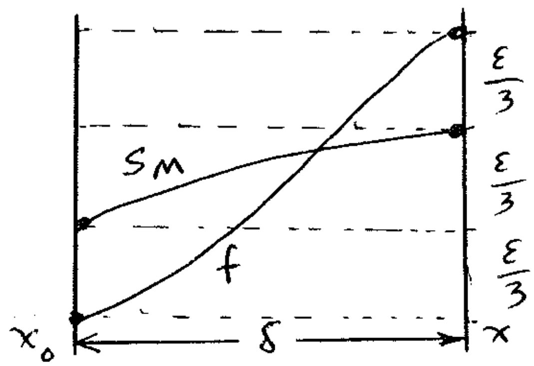
Proof (THE FAMOUS $\epsilon/3$ ARGUMENT)*:* Choose M so large that

$$\max_{a \le x \le b} \left| f(x) - S_M(x) \right| < \frac{\epsilon}{3} .$$

Then choose δ so small that

$$\left| S_M(x) - S_M(x_0) \right| < \frac{\epsilon}{3} \quad \text{whenever } |x - x_0| < \delta.$$

Then for $|x - x_0| < \delta$, $|f(x) - f(x_0)|$ is less than ϵ. This shows that f is continuous at x_0 . Here is a sketch of the worst possible case:



Now we will prove two theorems about uniform convergence of Fourier series.

Lemma 1 (Weierstrass M-test): If $|w_n(x)| \le M_n$ for all $x \in [a, b]$, and $\sum_{n=0}^{\infty} M_n$ converges, then $\sum_{n=0}^{\infty} w_n(x)$ converges uniformly on $[a, b]$.

The proof of this lemma is a slight extension of the proof of the "comparison test" in the series chapter of your calculus textbook.

Uniform Convergence Theorem 1: If the numerical series

$$\sum_{n=1}^{\infty} \left(|a_n| + |b_n| \right)$$

converges, then the Fourier series

$$\sum_{n=0}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$$

converges uniformly on $[-\pi, \pi]$ (and hence on the whole real line) to some function $f(x)$ of period 2π.

Proof: $|a_n \cos(nx) + b_n \sin(nx)| \le |a_n| + |b_n|$. Apply Lemma 1.

Lemma 2 (Bessel's Inequality): Let $\phi(x)$ have Fourier coefficients α_n and β_n. Suppose that $\int_{-\pi}^{\pi} |\phi(x)|^2\, dx < \infty$ (which will certainly be true if ϕ is piecewise continuous, hence bounded). Then

$$2\alpha_0^2 + \sum_{n=1}^{\infty}(|\alpha_n|^2 + |\beta_n|^2) \le \frac{1}{\pi}\int_{-\pi}^{\pi} |\phi(x)|^2\, dx < \infty.$$

The proof of this lemma is postponed until we discuss "convergence in the mean".

Uniform Convergence Theorem 2: If f (a function of period 2π) is *continuous* and also *piecewise smooth*, then its Fourier series converges uniformly — to f. [Recall that the graph of such a function may have corners, but no jumps.]

Proof: Since f is continuous, we can integrate by parts this way:

$$\begin{aligned} b_n &\equiv \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\, dx \\ &= -\frac{1}{\pi n}\big[f(x)\cos(nx)\big]_{x=-\pi}^{x=\pi} + \frac{1}{\pi n}\int_{-\pi}^{\pi} f'(x)\,\cos(nx)\, dx. \end{aligned}$$

(We'll come back later to see why continuity of f is crucial here, and why continuity of f' isn't.) The endpoint term is 0 because of periodicity and continuity. To study the other term, we let f' be the ϕ in Lemma 2, so that we get $b_n = \alpha_n/n$. A similar calculation yields $a_n = -\beta_n/n$ for $n \ne 0$. Now

$$0 \le \left(|\alpha_n| - \frac{1}{n}\right)^2 = {\alpha_n}^2 - \frac{2}{n}|\alpha_n| + \frac{1}{n^2},$$

so

$$\frac{|\alpha_n|}{n} \le \frac{1}{2}\left({\alpha_n}^2 + \frac{1}{n^2}\right).$$

Similarly,

$$\frac{|\beta_n|}{n} \le \frac{1}{2}\left({\beta_n}^2 + \frac{1}{n^2}\right).$$

But

$$\sum_{n=1}^{\infty}\left({\alpha_n}^2 + {\beta_n}^2\right) \quad \text{converges (by Lemma 2, Bessel),}$$

and

$$\sum_{n=1}^{\infty}\frac{1}{n^2} \quad \text{converges (well known).}$$

Therefore

$$\sum_{n=1}^{\infty}\left(|a_n|+|b_n|\right) \quad \text{converges (comparison test).}$$

Therefore, the Fourier series converges uniformly, by Theorem 1.

To summarize the key ideas of this proof: (1) Piecewise smoothness implies that f' has Fourier coefficients that fall off at large n. (Bessel's inequality gives us a stronger conclusion in this direction than the Riemann–Lebesgue theorem did.) (2) Continuity allows an integration by parts that shows that the Fourier coefficients of f are (roughly speaking) $1/n$ times those of f'. This implies that the coefficients of f fall off *fast*, so that Theorem 1 applies.

Conversely, if the coefficients fall off fast enough, then f is continuous, because of Theorem 1 and the $\epsilon/3$ theorem. *It is a general principle in Fourier analysis that smoothness properties in "x-space" (such as continuity and differentiability) are correlated with various degrees of rapid falloff at infinity in "n-space".*

Here is an example that shows what goes wrong with our integration by parts when f is not continuous. Let

$$f(x) \equiv \begin{cases} 0 & \text{if } x<1, \\ 1 & \text{if } x>1. \end{cases}$$

Then $f'(x)=0$ except at $x=1$, where it is undefined. This tempts one to say that the integral of f' is defined and equals 0. We now see that

$$\int_{-\pi}^{\pi} f(x)\,dx = \pi - 1$$

but integration by parts gives

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\,dx &= x f(x)\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} x f'(x)\,dx \\ &= \pi - 0,\end{aligned}$$

which is the wrong answer. This paradox can be resolved by saying that f' includes a term proportional to the delta function at the jump (here $\delta(x-1)$). But then f' does not satisfy the hypotheses of Bessel's inequality, so the proof of Theorem 2 does not apply to it.

Corollaries: Like the pointwise convergence theorem, the (second) uniform convergence theorem has some obvious extensions.

1. The period can be arbitrary ($2L$ instead of 2π).

2. If f is defined on a finite interval, say $[-L, L]$, and satisfies

$$f(-L+0) = f(L-0)$$

in addition to being continuous and piecewise smooth, then its Fourier series converges uniformly (because f has a continuous periodic extension).

3. A similar theorem applies to the cosine series on $[0, L]$, with no extra restrictions at the endpoints (since the *even* periodic extension is automatically continuous at the endpoints).

4. For the sine series there is a similar theorem, but this time one needs

$$f(0) = 0 = f(L)$$

to have a continuous *odd* extension.

A DIVERGENT FOURIER SERIES

So far our convergence theorems have assumed that f is piecewise smooth (hence differentiable except possibly at isolated points). This assumption is actually too strong (not necessary for convergence). But continuity by itself is not sufficient for convergence. That is, there are periodic functions that are *continuous everywhere*, but whose Fourier series fail to converge at certain points. (The functions are very wiggly near those points — having infinitely many maxima and minima.)

Example: The function

$$f(x) = \sum_{k=1}^{\infty} \sum_{m=1}^{2^{k^2}} \frac{1}{k^2 m} \left[\cos\left(2^{k^2+1} - m\right) x - \cos\left(2^{k^2+1} + m\right) x\right]$$

has a Fourier series that diverges at $x = 0$.

References: R. L. Jeffery, *Trigonometric Series*, pp. 20–23; A. Zygmund, *Trigonometric Series*, pp. 298–301.

As you can guess from this example, bad functions like this are hard to find. If you can write down a simple formula for a bounded, continuous function, its series is probably OK.

To understand this example, recall that "not converging" doesn't necessarily mean going to infinity. For example, the numerical series

$$\sum_{n=0}^{\infty} (-1)^n = 1 - 1 + 1 - 1 + 1 - 1 + \cdots$$

is divergent. Its partial sums $S_M = \sum_{n=0}^{M} (-1)^n$ are

$$S_0 = 1, \quad S_1 = 0, \quad S_2 = 1, \quad S_3 = 0, \ldots.$$

If we group the terms in a special way, we may get a limit — but this result is not unique: If we look only at the even partial sums, we would conclude that the series converges to 1, but if we look only at the odd ones, we would get the answer 0.

In the Fourier series example, when $x = 0$ each inner sum (over m) contains two blocks of terms, first positive and then negative. The signs of the terms go something like this:

$$+ + - - \Big| + + + + - - - - \Big| + \cdots$$

If we add up all the cosines one by one, the partial sums oscillate forever; this is what it means to say that that Fourier series diverges. But if we add up by blocks, as the double sum directs us, then the signs compensate within each block and the sum of blocks (sum over k) converges. That is, a certain *subsequence* of the partial sums of the divergent Fourier series converges (uniformly) to a continuous function $f(x)$, which has the original series as its Fourier series. (For proof of these claims see the references.)

Summability

There are various ways of defining the sum of a series that doesn't literally converge. They do not always agree; we are *defining* the sum, not discovering a preexistent fact. The definition that has proved most useful for Fourier series is called *Cesàro summability* or *convergence of arithmetic means.*

The Nth arithmetic mean is the average of the first N partial sums:

$$\sigma_N \equiv \frac{1}{N} \sum_{M=0}^{N-1} S_M \,.$$

For the numerical example $\sum_{n=0}^{\infty}(-1)^n$ we have

$$\sigma_N = \frac{1+0+1+0+\cdots}{N} \to \frac{1}{2}.$$

Fejér's Theorem: If f is piecewise continuous, the arithmetic means of its Fourier series converge (even if its partial sums don't).

You will be guided through the construction of a proof of this theorem in a homework assignment. A fuller statement of the theorem will appear there.

Even if the Fourier series does converge, the sequence of arithmetic means may converge faster, and therefore the Cesàro summation method will be of practical value in numerical work. This often happens when the coefficients a_n and b_n fall off slowly at infinity (say like $1/n$). The partial sums oscillate, but the averaging process smooths the oscillations out. On the other hand, if the coefficients fall off fast (like $1/n^4$, or e^{-n}), or if you know that all the terms are positive, then the process of taking arithmetic means will slow the convergence down, so it should be avoided.

Convergence in the mean

The "mean" referred to here is an averaging over x, not over M as in Cesàro summation. In modern mathematical writing, this kind of convergence is more often called "convergence in L^2" or "convergence in Hilbert space". It is the most important type of convergence for quantum mechanics.

To obtain uniform convergence we had to consider a smaller space of "nice" functions than the one where pointwise convergence holds. This time we will move in the other direction, working with a large space of functions (called $L^2(-\pi,\pi)$) that includes the piecewise continuous functions on $[-\pi,\pi]$ and many others besides — including some that are unbounded.

Definitions: The function f is *square-integrable* over the domain $[-\pi,\pi]$ if

$$\|f\|^2 \equiv \int_{-\pi}^{\pi} |f(x)|^2\, dx < \infty.$$

(One also says "f is in L^2".) The number $\|f\|$ is the *norm* of f.

Example: The function $f(x) = |x|^{-\epsilon}$ is square-integrable if $\epsilon < \frac{1}{2}$. Note that if $\epsilon > 0$, this function is not piecewise continuous by our definition, because it approaches ∞ (is unbounded) as x approaches 0. The function is not defined *at* 0, but that does not affect the existence of the integral.

Theorem (Cauchy–Schwarz inequality): If f and g are square-integrable, then

$$\int_{-\pi}^{\pi} f(x)^*\, g(x)\, dx \equiv \langle f, g\rangle \le \|f\|\|g\| < \infty.$$

(The $*$ is included to allow complex-valued functions. It can be ignored if f is real.)

A measure of the "distance" between two functions (real-valued ones, for notational simplicity) is

$$\|f-g\|^2 = \|f\|^2 + \|g\|^2 - 2\langle f, g\rangle.$$

Written out, this is

$$\int_{-\pi}^{\pi} [f(x) - g(x)]^2\, dx = \int_{-\pi}^{\pi} f(x)^2\, dx + \int_{-\pi}^{\pi} g(x)^2\, dx - 2\int_{-\pi}^{\pi} f(x)g(x)\, dx.$$

Suppose f is square-integrable and we want to approximate it by a *finite* sum of sines and cosines:

$$f(x) \approx A_0 + \sum_{n=1}^{M} \big(A_m \cos(nx) + B_n \sin(nx)\big) \equiv g_M(x).$$

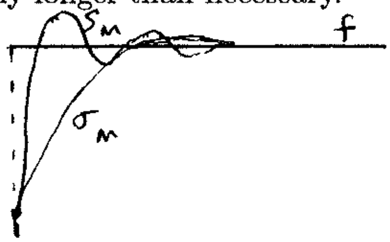
Which approximation is the best? Well, that depends on what you mean by "best"; a reasonable definition is "closest to f in the distance defined by the norm". That is, we pose the problem of minimizing

$$E_M \equiv \|f - g_M\|^2$$

(keeping M fixed, but varying the As and Bs). The E stands for "error".

One finds (see the textbooks for details) that the best g_M is the one for which $A_n = a_n$ and $B_n = b_n$ — that is, $g_M = S_M$, the truncated Fourier series of f.

In particular, σ_M is a worse approximation in the mean than S_M, even though in some cases it may be a better approximation at most points. This gives a new appreciation of the Gibbs phenomenon: It is the price S_M pays for not staying away from the limit function any longer than necessary.



The actual minimum value of E_M (attained when $g_M = S_M$) turns out to be

$$\min E_M = \|f - S_M\|^2 = \|f\|^2 - \pi\left[2{a_0}^2 + \sum_{n=1}^{M}({a_n}^2 + {b_n}^2)\right].$$

Since this number can't be negative, we see that

$$\|S_M\|^2 = \pi\left[2{a_0}^2 + \sum_{n=1}^{M}({a_n}^2 + {b_n}^2)\right] \le \|f\|^2 = \int_{-\pi}^{\pi} f(x)^2\,dx.$$

(If the function is complex, we should put absolute value bars around $f(x)$ and each of the as and bs here.) This is *Bessel's inequality*, already stated as Lemma 2 in the discussion of uniform convergence.

Now as $M \to \infty$ there are two logical possibilities:

1. $\min E_M \to 0$. Then we get **Parseval's equation**:

$$\boxed{2{a_0}^2 + \sum_{n=1}^{\infty}({a_n}^2 + {b_n}^2) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)^2\,dx.}$$

   In this case we have (by definition of E_M) that $\lim_{M\to\infty} \|f - S_M\| = 0$. That is, the series converges to f in the sense that the distance between the partial sum and f goes to 0. (Recall that, written out, this distance is the square root of a certain integral.) In this situation one says that the series converges "in the mean" or "in the topology of Hilbert space".

2. $\min E_M$ approaches a limit greater than 0. In this case S_M does not converge to f; there is a persistent error.

Although we can't prove it here, for Fourier series Alternative 1 holds:

Mean convergence theorem: If f is square-integrable, then its Fourier series converges to f in the mean, and Parseval's equation holds.

Notice that Alternative 2 would occur if we left some terms out of the series systematically. Suppose we didn't know anything about cosines and tried to write a Fourier series on the entire interval $[-\pi, \pi]$ containing only sines:

$$f(x) \stackrel{?}{=} \sum_{n=1}^{\infty} b_n \sin(nx).$$

Then whenever we had a function f that was not odd, our series would converge to the *odd part* of f,

$$\frac{f(x) - f(-x)}{2} = \sum_{n=1}^{\infty} b_n \sin(nx).$$

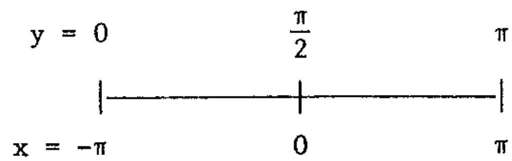
The error estimate would converge to the norm of the missing even part:

$$\begin{aligned} \lim_{M\to\infty} \left\| f - \sum_{n=1}^{M} b_n \sin(nx) \right\|^2 &= \left\| \sum_{n=1}^{\infty} a_n \cos(nx) \right\|^2 \\ &= \int_{-\pi}^{\pi} \left| \frac{f(x) + f(-x)}{2} \right|^2 dx. \end{aligned}$$

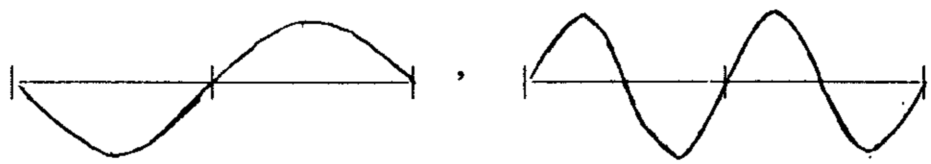
The failure of the series to converge is due not to a pathology of the function, but to the fact that our set of basic functions is *incomplete.*

[Please interpolate here two pages of old, typewritten notes.]

Consider a line segment, equipped with two coordinate systems related by $x = 2y - \pi$:



The functions $\{\sin nx\}$ $(n = 1, 2, 3, \ldots)$ are the same, up to sign, as the functions $\{\sin my\}$ $(m = 2, 4, 6, \ldots)$. Their graphs look like

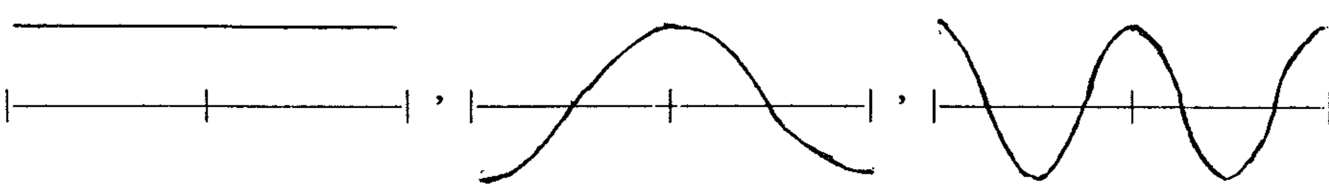


, etc.

$\sin x = -\sin 2y$ $\qquad$ $\sin 2x = \sin 4y$

They constitute an <u>incomplete</u> set, because the only functions which can be expressed in terms of them are the functions which are odd with respect to reflection about the midpoint of the line segment (i.e., odd in the variable x).

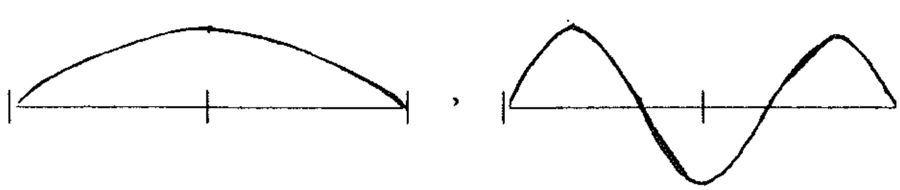
We can form a <u>complete</u> set by adding the functions $\{\cos nx\}$ $(n = 0, 1, 2, \ldots)$. Their graphs look like



$\cos 0x = 1$ $\qquad$ $\cos x$ $\qquad$ $\cos 2x$

etc. Every (square-integrable) function on the line segment can be expanded in terms of these functions (the two sets together). This is the ordinary Fourier series in the variable x. Its extension to the whole line represents a function with period equal to the length of the line segment.

Alternatively, we can form a <u>complete</u> set by adding to the set $\{\sin my\}$ (m even) the functions $\{\sin my\}$ $(m = 1, 3, 5, \ldots)$. These are the same, up to sign, as the functions $\{\cos (mx/2)\}$ $(m = 1, 3, 5, \ldots)$. Their graphs look like



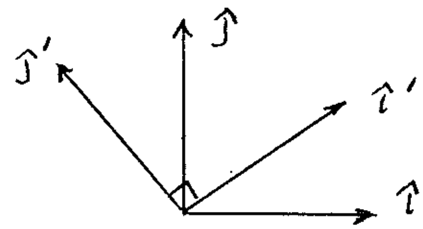
, etc.

$\cos \frac{x}{2} = \sin y$ $\qquad$ $-\cos \frac{3x}{2} = \sin 3y$

Every function on the line segment can be expanded in terms of these functions. This is the Fourier sine series in the variable y. Its extension to the whole line represents a function, with period equal to <u>twice</u> the length of the line segment, which is odd with respect to reflection about the endpoints of the segment (odd in the variable y, and also, as it happens, in the variable $u = \pi - y$).

In each case the added set of functions served as a basis for expanding all the functions that are even in x (i.e., even with respect to the midpoint of the segment). <u>Geometrical analogy</u>: A basis for expanding all vectors in the x-y plane is provided by the unit vectors $\hat{\imath}$ and $\hat{\jmath}$. An alternative basis consists of the vectors

$$\hat{\imath}' = \frac{1}{\sqrt{2}}(\hat{\imath} + \hat{\jmath}) \qquad \text{and} \qquad \hat{\jmath}' = \frac{1}{\sqrt{2}}(\hat{\jmath} - \hat{\imath}) \; .$$



Either set can be added to $\{\hat{k}\}$ (the set consisting of the unit vector in the z direction) to form a basis for expanding all vectors in three-dimensional space.

Fourier analysis as linear algebra
(functions as vectors)

I hope to convince you that the footnote on p. 47 of Powers is wrong: the word "not" should be deleted!

Look at these analogies between $L^2(-\pi, \pi)$ (the space of square-integrable functions on the domain $-\pi < x < \pi$) and $\mathbf{R}^3$ (the familiar space of three-dimensional vectors).

Concept	$\mathbf{R}^3$	L^2
vector	point, $\vec{x} = (x_1, x_2, x_3)$ index, $j = 1, 2, 3$	function, $f(x)$ variable, x
norm (length)	$\lvert\vec{x}\rvert = \sqrt{x_1{}^2 + x_2{}^2 + x_3{}^2}$	$\lVert f\rVert = \sqrt{\int_{-\pi}^{\pi} f(x)^2\, dx}$
inner product (scalar product)	$\vec{x}\cdot\vec{y} = \sum_{j=1}^{3} x_j y_j$ $\lVert\vec{x}\rVert^2 = \vec{x}\cdot\vec{x}$	$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x) g(x)\, dx$ $\lVert f\rVert^2 = \langle f, f\rangle$
orthogonal set	$\{\vec{e}_i\}$, where $\vec{e}_i\cdot\vec{e}_j = 0$ if $i \neq j$	$\{\phi_i(x)\}$, where $\langle\phi_i, \phi_j\rangle = 0$ if $i \neq j$ *Example:* $\{\cos(nx)\}$
complete set	Every $\vec{x} = \sum_{i=1}^{3} c_i \vec{e}_i$ (c_i are numbers)	Every $f(x) = \sum_{i=1}^{\infty} c_i \phi_i(x)$, convergence in mean

Remark: If the $\vec{e}_i$ or ϕ_i are orthogonal, they are linearly independent and hence the c_i for a given vector are unique. In fact,

$$c_i = \frac{1}{\lVert\phi_i\rVert^2}\langle\phi_i, f\rangle.$$

A set that is both linearly independent and complete (spanning) is called a *basis* for the vector space involved.

dimension	Every basis contains exactly three vectors $\vec{e}_i$.	Every basis contains infinitely many functions ϕ_i.
Parseval equation (for a basis!)	$\lvert\vec{x}\rvert^2 = \sum_{i=1}^{3} c_i^2 \lvert\vec{e}_i\rvert^2$	$\lVert f\rVert^2 = \sum_{i=1}^{\infty} c_i^2 \lVert\vec{f}_i\rVert^2$

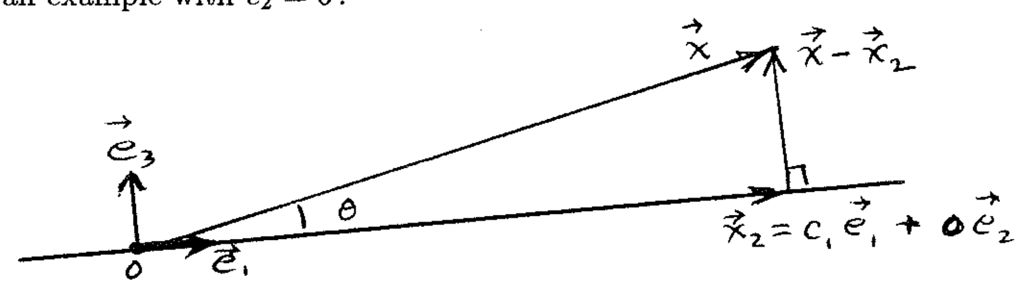
Remark: One often simplifies these equations by choosing $\lvert\vec{e}_i\rvert = 1$ or $\lVert\phi_i\rVert = 1$. However, in our notation for Fourier series we have

$$\lVert\cos(nx)\rVert^2 = \int_{-\pi}^{\pi} \cos^2(nx)\, dx = \pi = \lVert\sin(nx)\rVert^2,$$

$$\|\cos(0x)\|^2 = \int_{-\pi}^{\pi} 1\,dx = 2\pi.$$

Bessel inequality	If the set is orthogonal but not complete, then the best approximation to	
	$\vec{x}$	f
	of the form	
	$\vec{x}_N = \sum_{i=1}^{N} c_i\vec{e}_i \ (N = 1, 2)$	$g_N = \sum_{i=1}^{N \text{ (or } 2N+1)} c_i\phi_i(x)$
	— i.e., the one that minimizes the distance	
	$\lvert\vec{x} - \vec{x}_N\rvert$	$\|f - g_N\|$
	— satisfies	
	$\lvert\vec{x}_N\rvert^2 = \sum_{i=1}^{N} c_i^2\lvert\vec{e}_i\rvert^2 \le \lvert\vec{x}\rvert^2$	$\|g_N\|^2 = \sum_{i=1}^{N} c_i^2\|\phi_i\|^2 \le \|f\|^2$

Remark: If $\vec{e}_1$ and $\vec{e}_2$ span a plane, then $\vec{x}_2 = \sum_{i=1}^{2} c_i\vec{e}_i$ is the point in that plane which is closest to $\vec{x}$. This point is the intersection of the plane with the perpendicular dropped there from $\vec{x}$. The vector connecting $\vec{x}_2$ to $\vec{x}$ is $\vec{x} - \vec{x}_2 = c_3\vec{e}_3$. Here is an example with $c_2 = 0$:



Similarly, the sine terms of the Fourier series of f give the closest approximation to f by an odd function, and the remainder is an even function orthogonal to it.

Coefficients of the closest approximation	$c_i = \dfrac{\vec{x}\cdot\vec{e}_i}{\lvert\vec{e}_i\rvert^2} = \dfrac{\lvert\vec{x}\rvert\cos\theta}{\lvert\vec{e}_i\rvert}$	$c_i = \dfrac{\langle\phi_i, f\rangle}{\|\phi_i\|^2} = \dfrac{\int_{-\pi}^{\pi} f(x)\phi_i(x)\,dx}{\int_{-\pi}^{\pi}\phi_i(x)^2\,dx}$

In the Fourier case, this becomes

$$a_n = \frac{\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx}{\int_{-\pi}^{\pi}\cos^2(nx)\,dx} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx,$$

etc.

Mathematics 312 Fall, 1993

SPECIAL ASSIGNMENT DUE FRIDAY, OCTOBER 29

Prove Fejér's Theorem:

Let $S_M(x)$ be the partial sums of the Fourier series of a piecewise continuous periodic function $f(x)$. Let $\sigma_N(x)$ be the arithmetic means,

$$\sigma_N(x) = \frac{1}{N}\left[S_0(x) + S_1(x) + \cdots + S_{N-1}(x)\right].$$

Then as $N \to \infty$, $\sigma_N(x)$ converges pointwise to $\frac{1}{2}[f(x+0) + f(x-0)]$. Furthermore, if f is *continuous*, then $\sigma_N(x)$ converges *uniformly* to $f(x)$.

The proof of this theorem is very similar to the proof of pointwise convergence of the $S_M(x)$ for a piecewise smooth f. Follow these steps:

1. Show that $\displaystyle\sum_{n=0}^{N-1} \sin\left[\left(n+\tfrac{1}{2}\right)u\right] = \frac{\sin^2\left(\frac{Nu}{2}\right)}{\sin\left(\frac{u}{2}\right)}$.

   Hint: Use $2\sin^2 x = 1 - \cos(2x)$.

2. Show that $\displaystyle\sigma_N(x) = \frac{1}{2\pi N}\int_{-\pi}^{\pi} f(x+u)\,\frac{\sin^2\left(\frac{Nu}{2}\right)}{\sin^2\left(\frac{u}{2}\right)}\,du.$

   (Yes, the denominator involves $\sin$ in (1) but $\sin^2$ in (2).)

3. Show that $\displaystyle\frac{1}{2\pi N}\int_{0}^{\pi} \frac{\sin^2\left(\frac{Nu}{2}\right)}{\sin^2\left(\frac{u}{2}\right)}\,du = \frac{1}{2}$.

   Hint: Set $f(x) = 1$ in (2).

4. Show that to prove the pointwise convergence it is sufficient to prove

$$\lim_{N\to\infty} \frac{1}{2\pi N}\int_{0}^{\pi} [f(x+u) - f(x+0)]\,\frac{\sin^2\left(\frac{Nu}{2}\right)}{\sin^2\left(\frac{u}{2}\right)}\,du = 0$$

   and a similar statement for $f(x-0)$.

5. Show that (given ϵ) if $\delta > 0$ is sufficiently small, then for all N

$$I_1(x) \equiv \left|\frac{1}{2\pi n}\int_{0}^{\delta} [f(x+u) - f(x+0)]\,\frac{\sin^2\left(\frac{Nu}{2}\right)}{\sin^2\left(\frac{u}{2}\right)}\,du\right| < \frac{\epsilon}{2}.$$

6. Show that (given δ) if N is sufficiently large, then

$$I_2(x) \equiv \left| \frac{1}{2\pi n} \int_{\delta}^{\pi} [f(x+u) - f(x+0)] \frac{\sin^2\left(\frac{Nu}{2}\right)}{\sin^2\left(\frac{u}{2}\right)} \, du \right| < \frac{\epsilon}{2}.$$

7. Show that if f is continuous everywhere, the δ and N in (5) and (6) can be chosen to be the same for all x, and the convergence of the σ_N is therefore uniform. *Hint:* Use "Lemma 6" from the proof of the Riemann–Lebesgue theorem.

Appendix C

Commentary on the Michelson–Love–Gibbs Letters (**Nature**, 1898–1899)

1. Michelson, October 6

Michelson was one of the most distinguished experimental physicists of his era, now most famous for the Michelson–Morley experiment that later was interpreted as evidence for special relativity.

Paragraphs 3–5: Michelson quotes from Byerly's book the Fourier series of a sawtooth wave and a description of its graph. The function has a jump discontinuity at $x = \pi$. For clarity Michelson makes a change of variables so that the discontinuity appears at the origin, $\epsilon = 0$.

Paragraph 6: He argues that for a fixed, but arbitrarily large, n, ϵ can be chosen so small that all n terms can be approximated by the first terms of their power series. Thus each term is equal to ϵ, and the partial sum is equal to $n\epsilon$ (approximately).

Paragraph 7: He concludes, correctly, that the graph of the partial sum near the point of discontinuity is approximately a steep straight line (with slope opposite to that of the main part of the graph). However, either Michelson or the typesetter made a mistake here: The equation $-y = nx$ surely should be $-y = n(x - \pi)$ (or $-y = n\epsilon$).

Paragraph 9: This is a continuation of Paragraph 8: in that paragraph he has concluded that some function (actually, the nth partial sum) has a derivative at $\epsilon = 0$, and hence a tangent line.

Michelson is failing to distinguish between the pointwise convergent infinite sum and its partial sums. This is a major source of the confusion in the next two letters.

2. Love, October 13

Love later wrote a famous treatise on hydrodynamics, so he was not unfamiliar with "applied" mathematics.

Love has difficulty understanding the details of Michelson's remarks, so instead of rebutting them individually, he quotes them verbatim (including the typo!) and presents the standard, correct mathematical description of the convergence of the series, followed by a list of the common errors in reasoning that he presumes are infecting Michelson's thinking.

Alas, they are still talking past each other.

3. Michelson, December 29

Paragraph 1: Michelson wants to study a particular partial sum (a fixed n) and investigate its behavior when ϵ (now renamed x) is close to zero. Love wants to consider a fixed, nonzero x and take n to infinity, obtaining the sum of the infinite series. What Michelson says is correct for the partial sum, but irrelevant to the sum of the whole series.

Paragraph 2: For the partial sum, the "sawtooth" description is accurate only for $x > \pi/n$. Michelson therefore sees Love's remarks as irrelevant to Michelson's point.

Michelson's problem seems to be that he has what a historian of philosophy would call "a horror of the actual infinite". He has no concept of the sum of the infinite series as a single, finished, mathematical object; he thinks of the series as just the sequence of its partial sums. To put it more positively, although he expresses it in a confused way, he has an intuitive understanding of the crucial point that the limits

$$\lim_{x\to 0}\lim_{n\to\infty} \qquad \text{and} \qquad \lim_{n\to\infty}\lim_{x\to 0}$$

are not the same thing. For some reason he thinks that only the second limit is important; of course, the theory of convergence of Fourier series is about the first one.

4. Gibbs, December 29

Gibbs attempts to mediate the dispute. The last paragraph of his second letter is an excellent summary of this first letter: Love is talking about the *graph of the limit* of the partial sums, while Michelson is talking about the *limit of the graphs* of the partial sums.

5. Love, December 29

Love gives a better explanation (than in his first letter) of the motivation and conceptual background of the standard mathematical treatment of the convergence of infinite series of functions. It is a theory about *functions*, not about *curves*;

curves arise only for purposes of illustration. (And how much easier these letters would have been to understand, if *Nature* had been able and willing to print some!) For that reason, Love considers the graph of the limit to be the worthwhile object of discussion, and the limit of the graph to be relatively uninteresting.

6. Gibbs, April 27

The final paragraph has been discussed above. In the earlier paragraphs Gibbs announces his [re]discovery of the "Gibbs phenomenon", described in the context of the "limit of the graph" concept.

The Gibbs phenomenon was actually discovered in 1848 by someone named Wilbraham, who was looking at the square wave instead of the sawtooth. The history is described by H. S. Carslaw, *Bull. Amer. Math. Soc.* **31**, 420–424 (1925).

J. W. Gibbs (the "Willard" was not part of his surname, as Love apparently thought) could fairly be described as a physicist, mathematician, or chemist. His main accomplishments were (1) building the foundations of thermodynamics and statistical mechanics as applied to physical chemistry, and (2) formulating 3-dimensional vector calculus in essentially the way it is taught to and used by physicists and engineers today. The Gibbs phenomenon was a minor episode in his career.