CONVERGENCE OF FOURIER SERIES

0. Basic Information: This is just a collection of things we will need to develop the convergence theory, all in one place for handy reference. We'll do everything on the interval $[-\pi, \pi]$, because it is easy to "rescale" for more general intervals. Every function that we encounter on $[-\pi, \pi]$ will be assumed to be defined on the entire line $\mathbb{R}$ and to be periodic with period 2π, since it is easy to extend functions that were originally defined only on $[-\pi, \pi]$—but *be careful about endpoints*, since discontinuity of the function and/or its derivatives is easily introduced by uncritical redefinition at endpoints.

The reader should notice that other than the fact that we define the inner product of functions by an integral, we work exclusively with algebraic properties of the inner product, namely:

(1) The inner product is defined for every pair of (Riemann-integrable) functions f and g;

(2) $\langle f, f\rangle \geq 0$ for every f, and $\langle f, f\rangle = 0$ holds if and only if $f \equiv 0$ (except perhaps at a "small"—usually finite—set of points);

(3) The inner product is (conjugate-)symmetric: for every f and g, $\langle f, g\rangle = \overline{\langle g, f\rangle}$ (the complex conjugation is vacuous for real f and g);

(4) The inner product is linear in its first argument: $\langle \alpha_1 f_1 + \alpha_2 f_2, g\rangle = \alpha_1\langle f_1, g\rangle + \alpha_2\langle f_2, g\rangle$ holds for any (complex) scalars α_1, α_2 and functions f_1, f_2, g. (Of course, it is then *conjugate*-linear in its second argument: $\langle f, \alpha_1 g_1 + \alpha_2 g_2\rangle = \overline{\alpha}_1\langle f, g_1\rangle + \overline{\alpha}_2\langle f, g_2\rangle$.)

(5) The "length" or **norm** of vectors, defined by $\|f\| = \sqrt{\langle f, f\rangle}$, has the properties $\|\alpha f\| = |\alpha|\,\|f\|$ and $\|f + g\| \leq \|f\| + \|g\|$.

Everything we list in this § is thus equally valid for the finite-dimensional vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, because their usual inner or "dot" product shares these algebraic properties.

0.1. The Inner Product of Functions: The **inner** or dot **product** is **defined** by[(1)]

$$\langle f, g\rangle = \begin{cases} \displaystyle\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx & \text{when considering trigonometric series;} \\ \displaystyle\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx & \text{when considering complex-exponential series.} \end{cases}$$

The complex-conjugation bar is vacuous when one deals with real-valued functions. Note that because of our assumption of periodicity, the integration over $[-\pi, \pi]$ can be replaced by an integration over any interval $[a - \pi, a + \pi]$ without changing the value of the integral.

Just as in $\mathbb{R}^n$ or $\mathbb{C}^n$, we say that two functions f and g are **orthogonal** (or, somewhat loosely, **perpendicular**) if their inner product $\langle f, g\rangle = 0$. This relation is occasionally written as $f \perp g$.

The L^2 or **root-mean-square** (= **r. m. s.**) **norm** of a function is **defined** (omitting the factor $\dfrac{1}{2\pi}$ when appropriate) by

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

There is a "law of cosines" for this distance relation in view of the fact that

$$\begin{aligned} \|f + g\|^2 &= \langle f + g, f + g\rangle \\ &= \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle \\ &= \begin{cases} \|f\|^2 + 2\mathrm{Re}\,[\langle f, g\rangle] + \|g\|^2 & \text{(complex scalars);} \\ \|f\|^2 + 2\,\langle f, g\rangle + \|g\|^2 & \text{(real scalars).} \end{cases} \end{aligned}$$

(1) All these definitions make sense for improper Riemann integrals; but it is so easy to do all this stuff when Lebesgue integration theory is employed that there is not much reason to work through the details of extending these definitions here.

So the "Pythagorean theorem" holds for orthogonal vectors: if $f \perp g$ (so $\langle f, g\rangle = 0$) then

$$\|f+g\|^2 = \|f\|^2 + \|g\|^2 \ .$$

The (Cauchy-Buniâkovskiĭ-)**Schwarz inequality** controls the size of an inner product by the sizes of the factors:

$$|\langle f, g\rangle| \le \|f\| \cdot \|g\| \ .$$

It can be proved as follows. Write $\langle f, g\rangle$ in the "complex polar" form $\langle f, g\rangle = re^{i\theta}$, where $r = |\langle f, g\rangle|$, and let $h = e^{i\theta}g$; then $\langle f, h\rangle = e^{-i\theta}re^{i\theta} = r \ge 0$. For real t consider the everywhere-nonnegative function of t given by the expression {note that $\langle h, h\rangle = \langle g, g\rangle = \|g\|^2$}

$$\|f - th\|^2 = \langle f - th, f - th\rangle = \langle f, f\rangle - 2t\,\langle f, h\rangle + t^2\,\langle h, h\rangle = \|f\|^2 - 2rt + t^2\,\|g\|^2 \ge 0 \ .$$

If $\|g\| = 0$ then this expression can remain nonnegative for all $t \in \mathbb{R}$ if and only if $|\langle f, g\rangle| = r = 0$, establishing the Schwarz inequality in that case (with both sides equal to zero). If $\|g\| > 0$ then the graph of this expression is a parabola opening upward; we can find its vertex by computing the zero of its derivative and plugging in:

$$\begin{aligned}
\frac{d}{dt}\left[\|f\|^2 - 2rt + t^2\,\|g\|^2\right] &= -2r + 2t\,\|g\|^2 \\
t &= \frac{r}{\|g\|^2} \\
0 &\le \|f\|^2 - 2r\cdot\frac{r}{\|g\|^2} + \left(\frac{r}{\|g\|^2}\right)^2 \|g\|^2 = \|f\|^2 - r\cdot\frac{r}{\|g\|^2} \\
|\langle f, g\rangle|^2 = r^2 &\le \|f\|^2 \cdot \|g\|^2
\end{aligned}$$

which (after square roots of both sides are taken) is the Schwarz inequality.

The Schwarz inequality implies the **triangle inequality**: take the square root of the beginning and end of

$$\begin{aligned}
\|f+g\|^2 &= \langle f+g, f+g\rangle \\
&= \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle \\
&= \begin{cases} \|f\|^2 + 2\mathrm{Re}\,[\langle f, g\rangle] + \|g\|^2 & \text{(complex scalars);} \\ \|f\|^2 + 2\,\langle f, g\rangle + \|g\|^2 & \text{(real scalars)} \end{cases} \\
&\le \|f\|^2 + 2\,\|f\|\,\|g\| + \|g\|^2 = \Big(\|f\| + \|g\|\Big)^2 \quad \text{to get} \\
\|f+g\| &\le \|f\| + \|g\| \ .
\end{aligned}$$

A sequence (finite or infinite) $\{X_n\}$ of functions is **orthogonal** if $\langle X_m, X_n\rangle = 0$ for $m \ne n$. It is usual to consider only orthogonal sequences of functions that are "essentially nonzero," so that $\langle X_n, X_n\rangle > 0$ for each n. By replacing each X_n by $\phi_n = \dfrac{1}{\|X_n\|}X_n$ we obtain a orthogonal sequence $\{\phi_n\}$ of functions with the additional property that $\|\phi_n\| = 1$ for all n; such a sequence is called an **orthonormal** sequence.

0.2. The Bessel-Parseval Relations: These are what one gets by looking critically at the results of the computation on Strauss's pp. 127–128. Given a function f, an orthogonal set $\{X_n\}$ of functions and a corresponding sequence of constants $\{c_n\}$, we form a *finite* sum $\sum c_n X_n$ (the easiest way to indicate this is to take a finite subset J of the indices and write $\sum\limits_{n\in J}\cdots$) and compute its r. m. s.-distance-squared from f, obtaining (per Strauss,[2] p. 127)

[2] We have added complex-conjugation and absolute-value in a few cases to extend Strauss's real-case computations to the complex case.

$$\left\| f - \sum_{n\in J} c_n X_n \right\|^2 = \left\langle f - \sum_{n\in J} c_n X_n, f - \sum_{n\in J} c_n X_n \right\rangle$$
$$= \left\{ \sum_{n\in J} \|X_n\|^2 \cdot \left| c_n - \frac{\langle f, X_n\rangle}{\|X_n\|^2} \right|^2 \right\} + \left\{ \|f\|^2 - \sum_{n\in J} \frac{|\langle f, X_n\rangle|^2}{\|X_n\|^2} \right\} . \tag{15}$$

One now begins to read interesting facts off this identity.

(1) The expressions inside the large pairs of {} braces are nonnegative. Thus the only way to make $\left\| f - \sum_{n\in J} c_n X_n \right\|$ as small as possible (for a given set J of indices) is to choose each c_n to have the corresponding value $\dfrac{\langle f, X_n\rangle}{\|X_n\|^2}$ (any other choice will put a strictly-positive term into the sum inside the first large braces). If (and only if) these choices are made for all indices $n \in J$, we shall have

$$0 \le \left\| f - \sum_{n\in J} c_n X_n \right\|^2 = \|f\|^2 - \sum_{n\in J} \frac{|\langle f, X_n\rangle|^2}{\|X_n\|^2} . \tag{16}$$

This relation has a "geometrical" meaning. The sums of the form $\sum_{n\in J} c_n X_n$ are the elements of the vector subspace (of the space of all integrable functions) spanned by $\{X_n : n \in J\}$. Relations (15) and (16) imply that the element of that vector subspace that is closest to f in the L^2-norm is the sum of the projections on the 1-dimensional subspaces spanned by the individual X_n's. {Recall the sophomore-calculus formula

$$\mathrm{proj}_{\mathbf{x}}\mathbf{f} = \frac{\mathbf{f} \bullet \mathbf{x}}{\|\mathbf{x}\|^2}\, \mathbf{x}$$

for the (perpendicular) projection of a vector **f** onto the (line generated by a) vector **x**.} This is a generalization of the fact that, for example, the point of the **ij**-plane in $\mathbb{R}^3$ that is closest to the vector $a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$ is $a_1\,\mathbf{i} + a_2\,\mathbf{j}$.

It's easy to see geometrically that the closest point is also the "foot of the perpendicular" dropped from the given vector to the subspace, and the same is true in function space:

(2) The sum $\sum_{n\in J} c_n X_n$ that is closest to f in norm is characterized by the relation

$$f - \sum_{n\in J} c_n X_n \perp X_m \quad \text{for every } m \in J.$$

Indeed, for every $m \in J$

$$f - \sum_{n\in J} c_n X_n \perp X_m \iff \langle f, X_m\rangle - \sum_{n\in J} c_n \langle X_n, X_m\rangle = 0$$
$$\iff \langle f, X_m\rangle - c_m \langle X_m, X_m\rangle = 0 \iff c_m = \frac{\langle f, X_m\rangle}{\|X_m\|^2} .$$

These coefficients $c_n = \dfrac{\langle f, X_n\rangle}{\|X_n\|^2}$ are called the **components** or the **abstract Fourier coefficients** of f (**with respect to** $\{X_n\}$).

The rather surprising deduction from (16) is, that since

$$0 \leq \left\| f - \sum_{n \in J} c_n X_n \right\|^2 = \|f\|^2 - \sum_{n \in J} \frac{|\langle f, X_n \rangle|^2}{\|X_n\|^2}$$

can be rewritten as the **Bessel inequality**

$$\sum_{n \in J} \frac{|\langle f, X_n \rangle|^2}{\|X_n\|^2} \leq \|f\|^2 \ ,$$

it turns out to be true that

(3) The scalar series $\sum_n \frac{|\langle f, X_n \rangle|^2}{\|X_n\|^2}$ converges, no matter in what order the sequence $\{X_n\}$ is given. The sum of this series is $\leq \|f\|^2$. Moreover, in order that the series $\sum_n \frac{\langle f, X_n \rangle}{\|X_n\|^2} X_n$ should converge to f in the sense of the r. m. s. norm, it is necessary and sufficient that the scalar series $\sum_n \frac{|\langle f, X_n \rangle|^2}{\|X_n\|^2}$ should converge to $\|f\|^2$.

See the discussion on Strauss's p. 128. The relation $\sum_n \frac{|\langle f, X_n \rangle|^2}{\|X_n\|^2} = \|f\|^2$, when it holds (*i.e.*, when the series converges to f in r. m. s. norm), is called the **Parseval equation** or **Parseval relation**.

Before taking leave of this subject we should observe that the special case of the Bessel inequality in which there is only one X_n is the Schwarz inequality: replacing X_n by the single function g gives (with a one-term sum $\sum_{n \in J} \cdots$)

$$\frac{|\langle f, g \rangle|^2}{\|g\|^2} \leq \|f\|^2$$
$$|\langle f, g \rangle|^2 \leq \|f\|^2 \|g\|^2 \ .$$

0.3. Everybody's Favorite Orthogonal Sequences: These are the ones given by $\{\sin nx\}_{n=1}^{\infty}$ (typically on $[0, \pi]$), by $\{\cos nx\}_{n=0}^{\infty}$ (typically on $[0, \pi]$), by the union of those two sequences on $[-\pi, \pi]$, and[(3)] by $\{e^{inx}\}_{n \in \mathbb{Z}}$ on $[-\pi, \pi]$. The functions in the first of these sequences are the eigenfunctions of $-D^2$ with Dirichlet end conditions on $[0, \pi]$, the second are the eigenfunctions of $-D^2$ with Neumann end conditions on $[0, \pi]$, and their union makes up the eigenfunctions of $-D^2$ with periodic end conditions on $[-\pi, \pi]$. The complex exponentials are the eigenfunctions of $-iD$ on $[-\pi, \pi]$ with periodic end conditions. Thus Sturm-Liouville theory tells us automatically that these sequences are orthogonal (although that can be verified directly by concrete integration). The Euler relations $e^{\pm i\theta} = \cos\theta \pm i \sin\theta$ tell us that the 2-dimensional complex function space spanned by $\cos nx$ and $\sin nx$ is the same as the complex function space spanned by e^{inx} and e^{-inx}.

The "abstract Fourier coefficients" become the "concrete Fourier coefficients" for these sequences (for simplicity's sake we shall only consider the case of periodic end conditions on $[-\pi, \pi]$): if $X_n = \cos nx$ then

$$\|X_n\|^2 = \int_{-\pi}^{\pi} \cos^2 nx \, dx = \pi$$
$$A_n = \frac{\langle f, X_n \rangle}{\|X_n\|^2} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx$$

(3) Recall the standard "blackboard-bold" letter names customarily used for some of the standard sets of mathematics: $\mathbb{N}$ for the natural numbers, $\mathbb{Z}$ for the signed integers (whole numbers), $\mathbb{Q}$ for the rational numbers (quotients of integers), as well as the more familiar $\mathbb{R}$ for the reals and $\mathbb{C}$ for the complex numbers.

(the exceptional case $n = 0$ is conventionally handled by writing the $n = 0$ term of the Fourier series in the form $\frac{A_0}{2}$) and similarly for $\sin nx$.[(4)] The orthogonal sequence $\{X_n = e^{inx}\}_{n\in\mathbb{Z}}$ is particularly easy to work with, because it is orthonormal and no division by π is necessary (nor does the case $n = 0$ require any special treatment):

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

(note that the fact that X_n is the second argument of the inner product makes the corresponding factor in the integrand equal $\overline{e^{inx}} = e^{-inx}$). It is customary to "keep the n-th harmonics together" when writing (formal) Fourier series, either in real (trig function) or complex(-exponential) form, so that they appear as

$$f(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty}\{A_n \cos nx + B_n \sin nx\} \quad \text{or}$$
$$f(x) = \lim_{N\to\infty}\sum_{n=-N}^{N} c_n e^{inx}$$

respectively.[(5)] In either case, for $n > 0$ the term $\{A_n \cos nx + B_n \sin nx\}$ or $c_{-n}e^{-inx} + c_n e^{inx}$ is the only term that belongs to the linear subspace of function space spanned by $\cos nx$ and $\sin nx$, and so

$$A_n \cos nx + B_n \sin nx = (c_{-n} + c_n)\cos nx + i(-c_{-n} + c_n)\sin nx$$
$$A_n = c_n + c_{-n} \quad B_n = i(c_n - c_{-n})$$
$$\frac{A_n - iB_n}{2} = c_n \quad \frac{A_n + iB_n}{2} = c_{-n}\ .$$

Thus if $f(x)$ is real (making both A_n and B_n real), one must have $c_{-n} = \overline{c}_n$, and the converse implication also holds. (Obviously $\frac{A_0}{2} = c_0$, since it is the average value of $f(x)$ over $[-\pi, \pi]$ in both cases.)

One advantage possessed by the trig-series formulation is that trig functions respect **parity**: functions and their series split neatly into **even functions** (those satisfying $f(-x) \equiv f(x)$) and **odd functions** (those satisfying $f(-x) \equiv -f(x)$). The only (nonzero) terms in the Fourier expansion of an even function with period 2π are the cosine terms; the only terms in the expansion of an odd function with period 2π are the sine terms. For this reason, expansions of $f(x)$ in the eigenfunctions of $-D^2$ on an interval $[0, \pi]$ with Dirichlet or Neumann boundary conditions respectively can be handled by extending $f(x)$ to be odd and periodic with period 2π or even and periodic with period 2π respectively, then applying what we shall shortly find out[(6)] about Fourier expansions of periodic functions with period interval $[-\pi, \pi]$.

0.4. Translations: Questions of pointwise convergence of the Fourier series of a function $f(x)$ at arbitrary points $x \in [-\pi, \pi]$ can be reduced to questions of convergence at zero by examining the Fourier coefficients of the "translated" function f_{-x} defined by

$$f_{-x}(z) = f(z + x)\ .$$

---

(4) The letters a_n, b_n and c_n (or their upper-case or funny-font-style variants) are conventionally used to denote Fourier cosine, sine and complex-exponential coefficients respectively, and we shall adhere to that convention as much as is practical.

(5) The sign of equality between $f(x)$ on the one side and the sum of the series (or limit of the partial sums) on the other has to be interpreted carefully, since there are functions whose Fourier series do not converge to them pointwise. We will make a fuss over this below. Perhaps I should have put quotation marks around the sign of equality, or used a different symbol like "$\simeq$" or "$\rightleftharpoons$" or whatever.

(6) Unfortunately, general Sturm-Liouville expansions, like those in the eigenfunctions of Robin problems, require a different approach.

First the trig-function formulation (so that the reader can see how much simpler the complex-exponential formulation is). If we put

$$A_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f_{-x}(z)\cos nz\,dz$$

(retaining A_n to mean $A_n(0)$ and B_n to mean $B_n(0)$, for obvious reasons) then[(7)] the addition formula for the cosine gives us

$$\begin{aligned}
A_n(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f_{-x}(z)\cos nz\,dz = \frac{1}{\pi}\int_{-\pi}^{\pi} f(z+x)\cos nz\,dz \\
&= \frac{1}{\pi}\int_{-\pi+x}^{\pi+x} f(y)\cos n(y-x)\,dy = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\cos n(y-x)\,dy \quad \text{(everything has period } 2\pi\text{)} \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\cos ny\cos nx\,dy + \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\sin ny\sin nx\,dy = A_n\cos nx + B_n\sin nx\ .
\end{aligned}$$

Similarly,

$$B_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\sin ny\cos nx\,dy - \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\cos ny\sin nx\,dy = B_n\cos nx - A_n\sin nx\ .$$

The value of the n-th "complete" term of the trigonometric Fourier series of f_{-x} (including both the $\cos nz$ and $\sin nz$ terms) thus appears as

$$\begin{aligned}
A_n(x)\cos nz + B_n(x)\sin nz &= [A_n\cos nx + B_n\sin nx]\cos nz + [B_n\cos nx - A_n\sin nx]\sin nz \\
&= A_n[\cos nx\cos nz - \sin nx\sin nz] + B_n[\sin nx\cos nz + B_n\cos nx\sin nz] \\
&= A_n\cos n(z+x) + B_n\sin n(z+x)\ .
\end{aligned}$$

In words: the effect of plugging $z+x$ in for the argument in $f(z)$ is to change each complete n-th term of the Fourier series of f in exactly the same way as plugging $z+x$ in for the argument in $\cos nz$ and $\sin nz$ would have done.

The same argument for complex exponentials is easier: with $c_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(z)e^{-inz}\,dz$ (and with $c_n = c_n(0)$ in analogy to what we did before), we have

$$\begin{aligned}
c_n(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(z+x)e^{-inz}\,dz = \frac{1}{2\pi}\int_{-\pi+x}^{\pi+x} f(y)e^{-in(y-x)}\,dy \qquad (y = z+x) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}e^{inx}\,dz = c_n e^{inx} \qquad \text{(periodicity)} \\
c_n(x)e^{inz} &= c_n e^{inx}e^{inz} = c_n e^{in(z+x)}\ .
\end{aligned}$$

In either formulation, it is clear that investigating the convergence of the Fourier series of $f(z)$ to a value determined from $f(z)$ at $z = x$ is exactly the same problem as investigating the convergence of the Fourier series of $f_{-x}(z)$ to a value determined from $f_{-x}(z)$ at $z = 0$. **For this reason, we confine proofs of convergence of Fourier series to proofs of convergence at $z = 0$.**

0.5. The Dirichlet Kernel: It is possible to give a closed-form formula for the sum of the first terms of the complete Fourier series, corresponding to indices $0 \le n \le N$. We have aired the details in class and

---

(7) The coefficient A_0, since it is defined to be twice the average of the function over one period, does not change under translation—although one can also see that from the n=0 case of the formulas.

they are exposed on pp. 132–134 of Strauss's book, so we give only a synopsis. Writing the integrals that define the c_n's in the series gives

$$S_N(x) = \sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N} \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}\,dy\right] e^{inx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\left[\sum_{n=-N}^{N} e^{-iny}e^{inx}\right] dy$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\left[\sum_{n=-N}^{N} e^{-in(y-x)}\right] dy = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)K_N(x-y)\,dy$$

where $K_N(\theta) = \sum_{n=-N}^{N} e^{in\theta} = \dfrac{\sin\left(\dfrac{2N+1}{2}\,\theta\right)}{\sin\left(\dfrac{\theta}{2}\right)}$, the **Dirichlet kernel** (function).[8] A good portrait of $K_N(\theta)$ for fairly-good-sized N can be found on Strauss's p. 133. It is easy to see {*e.g.*, by term-by-term integration of the finite-sum expansion of $K_N(\theta)$} that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(\theta)\,d\theta = 1\ , \quad \text{and by even-ness (= symmetry), that therefore}$$
$$\frac{1}{2\pi}\int_{0}^{\pi} K_N(\theta)\,d\theta = \frac{1}{2}\ .$$

Unfortunately, it can be shown that the integral $\frac{1}{2\pi}\int_{-\pi}^{\pi} |K_N(\theta)|\,d\theta$ of the absolute value of $K_N(\theta)$ diverges to $+\infty$ as $N \to \infty$, and due to this fact there exist continuous functions with period 2π whose Fourier series fail to converge (let alone converge to the value of the function) at certain points. Thus proofs of convergence of Fourier series require rather stronger hypotheses like the Dini condition introduced below.

0.6. The Riemann-Lebesgue Lemma: See the separate notes on this subject for a proof of:

Let f be a Riemann-integrable function defined on an interval $a \le x \le b$ of the real line. Then for any real β

$$\lim_{|\alpha|\to\infty} \int_a^b f(x)\sin(\alpha x + \beta)\,dx = 0\ .$$

1. Convergence of the Fourier Series at Individual Points: The ingredients in the proof of the standard result on this subject are three: the formula that gives the Dirichlet kernel, the Riemann-Lebesgue lemma, and an additional hypothesis about the function $f(x)$ whose Fourier series is under consideration, namely

1.1 The Dini Condition: Let $f(x)$ be a function of period 2π defined on $\mathbb{R}$ and Riemann-integrable on $[-\pi,\pi]$ (and therefore on every finite interval). It satisfies the **Dini condition on the right** at a point $x \in [-\pi,\pi]$ if $f(x^+) = \lim_{t\to x^+} f(t)$ exists and there is some $\delta > 0$ such that the integral

$$\int_x^{x+\delta} \left|\frac{f(t)-f(x^+)}{t-x}\right| dt < +\infty\ .$$

The **Dini condition on the left** is defined similarly. Evidently if this integral is finite for some particular $\delta > 0$ it is finite for all smaller values of $\delta > 0$; moreover, because the integral in question (if it is improper) is defined by the limit

$$\int_x^{x+\delta} \left|\frac{f(t)-f(x^+)}{t-x}\right| dt = \lim_{\eta\to 0^+} \int_{x+\eta}^{x+\delta} \left|\frac{f(t)-f(x^+)}{t-x}\right| dt \quad \text{where } 0<\eta<\delta$$

[8] I apologize for my intellectual sloth in having called this function $D_N(\theta)$—its usual name—in class, rather than using Strauss's $K_N(\theta)$. I should have checked to see whether Strauss felt he couldn't call it by the same name that everybody else uses. Silly me.

it is clear that

$$\lim_{\delta\to 0^+}\int_x^{x+\delta}\left|\frac{f(t)-f(x^+)}{t-x}\right|\,dt = 0\ .$$

1.2 Theorem [of Ulisse Dini]: Let $f(x)$ be a function of period 2π defined on $\mathbb{R}$ and Riemann-integrable on $[-\pi,\pi]$ (and therefore on every finite interval). If it satisfies both the left and the right Dini conditions at a point x, then its Fourier series at x converges to $\dfrac{f(x^-)+f(x^+)}{2}$.

Lemma: The function $\dfrac{\psi}{\sin\psi}$ is an increasing function on the interval $0\le\psi\le\pi$ (we assume that the function has been extended to have the value 1 at $\psi=0$) and therefore bounded between 1 and $\dfrac{\pi}{2}$ on the interval $0\le\psi\le\pi/2$.

Proof of the lemma. The derivative of the function is $\dfrac{\sin\psi-\psi\cos\psi}{\sin^2\psi}$. The numerator is nonnegative for $0\le\psi<\dfrac{\pi}{2}$ in view of the fact that $\dfrac{\sin\psi}{\cos\psi}=\tan\psi\ge\psi$ on that interval (the tangent function has the increasing derivative $\sec^2\psi$, so its graph lies above its tangent line); thus the function increases from $\psi=0$ to $\psi=\dfrac{\pi}{2}$. On the other hand, for $\dfrac{\pi}{2}\le\psi\le\pi$, the numerator of $\dfrac{\psi}{\sin\psi}$ increases while the denominator decreases, so the quotient surely increases.

Proof of the theorem. In view of **0.4** above it suffices to consider the case $x=0$, and by the symmetry of the Dirichlet kernel it suffices to show that

$$\lim_{N\to\infty}\frac{1}{2\pi}\int_0^\pi f(\theta)K_N(\theta)\,d\theta=\frac{f(0^+)}{2}$$

when $f(x)$ satisfies the Dini condition on the right at $x=0$ (since the proof of the corresponding assertion for $f(x^-)$ is a mirror image of the one we will give here). Because $\dfrac{1}{2\pi}\displaystyle\int_0^\pi K_N(\theta)\,d\theta=\dfrac{1}{2}$, this is equivalent to showing that

$$\lim_{N\to\infty}\frac{1}{2\pi}\int_0^\pi\Big[f(\theta)-f(0^+)\Big]K_N(\theta)\,d\theta=0\ .$$

This is a formal limit proof, so let $\epsilon>0$ be given; we shall show that for every sufficiently large N the relation

$$\left|\frac{1}{2\pi}\int_0^\pi\Big[f(\theta)-f(0^+)\Big]K_N(\theta)\,d\theta\right|<\epsilon$$

holds. Begin by finding (and then holding fixed) $\delta>0$ for which

$$\frac{1}{2\pi}\int_0^\delta\left|\frac{f(\theta)-f(0^+)}{\theta}\right|\,d\theta<\frac{\epsilon}{2\pi}\ ;$$

then we can estimate $\left|\dfrac{1}{2\pi}\displaystyle\int_0^\delta\Big[f(\theta)-f(0^+)\Big]K_N(\theta)\,d\theta\right|$ by

$$\begin{aligned}
\left|\frac{1}{2\pi}\int_0^\delta\Big[f(\theta)-f(0^+)\Big]K_N(\theta)\,d\theta\right| &= \left|\frac{1}{2\pi}\int_0^\delta\Big[f(\theta)-f(0^+)\Big]\frac{\sin\left(\dfrac{2N+1}{2}\,\theta\right)}{\sin\left(\dfrac{\theta}{2}\right)}d\theta\right|\\
\text{(multiply and divide by }\theta/2\text{)}\qquad &= \left|\frac{1}{2\pi}\int_0^\delta 2\left[\frac{f(\theta)-f(0^+)}{\theta}\right]\frac{\dfrac{\theta}{2}}{\sin\left(\dfrac{\theta}{2}\right)}\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\right|\\
&\le\frac{1}{2\pi}\int_0^\delta 2\left|\frac{f(\theta)-f(0^+)}{\theta}\right|\cdot\frac{\pi}{2}\cdot 1\,d\theta<\frac{\epsilon}{2}\ ,
\end{aligned}$$

where we have used the immediately preceding lemma to estimate $\dfrac{\frac{\theta}{2}}{\sin\left(\frac{\theta}{2}\right)}$ and (of course) estimated $\left|\sin\left(\frac{2N+1}{2}\,\theta\right)\right|$ by 1. It is even simpler to handle the rest of the integral (over $[\delta,\pi]$), because in this interval $\sin\left(\frac{\theta}{2}\right) > \sin\left(\frac{\delta}{2}\right)$ holds. To see this, rewrite the integrand in the form shown below:

$$\frac{1}{2\pi}\int_{\delta}^{\pi}\left[\frac{f(\theta)-f(0^+)}{\sin\left(\frac{\theta}{2}\right)}\right]\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\ .$$

Now, since the function inside the square brackets [ ] is Riemann-integrable (the denominator is continuous and stays bounded away from zero in the interval of integration), we can apply the Riemann-Lebesgue lemma to see that this integral can be made arbitrarily small by taking N sufficiently large. If N_ϵ is such that $N \geq N_\epsilon \Longrightarrow \left|\frac{1}{2\pi}\int_{\delta}^{\pi}\left[\frac{f(\theta)-f(0^+)}{\sin\left(\frac{\theta}{2}\right)}\right]\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\right| < \frac{\epsilon}{2}$, then it is clear that also

$$\begin{aligned}
N \geq N_\epsilon \Longrightarrow & \left|\frac{1}{2\pi}\int_{0}^{\pi}\left[\frac{f(\theta)-f(0^+)}{\sin\left(\frac{\theta}{2}\right)}\right]\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\right| \\
&\leq \left|\frac{1}{2\pi}\int_{0}^{\delta}\left[\frac{f(\theta)-f(0^+)}{\sin\left(\frac{\theta}{2}\right)}\right]\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\right| + \left|\frac{1}{2\pi}\int_{\delta}^{\pi}\left[\frac{f(\theta)-f(0^+)}{\sin\left(\frac{\theta}{2}\right)}\right]\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\right| \\
&\leq \frac{1}{2\pi}\int_{0}^{\delta}2\left|\frac{f(\theta)-f(0^+)}{\theta}\right|\cdot\frac{\pi}{2}\cdot 1\,d\theta + \left|\frac{1}{2\pi}\int_{\delta}^{\pi}\left[\frac{f(\theta)-f(0^+)}{\sin\left(\frac{\theta}{2}\right)}\right]\sin\left(\frac{2N+1}{2}\,\theta\right)d\theta\right| < \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon
\end{aligned}$$

and that (together with its mirror image for $f(0^-)$) proves Dini's theorem.

The Dini condition may look strange, but it is very easy to give "reasonable-looking" properties of Riemann-integrable functions $f(x)$ that imply it. For example, if $f(x)$ has right- and left-hand limits at a point x and has left- and right-hand derivatives there, in the sense that the two limits

$$\lim_{\Delta x\to 0^+}\frac{f(x+\Delta x)-f(x^+)}{\Delta x} = f'^+(x) \quad\text{and}\quad \lim_{\Delta x\to 0^-}\frac{f(x+\Delta x)-f(x^-)}{\Delta x} = f'^-(x)$$

both exist, then the integrand $\frac{f(t)-f(x^+)}{t-x}$ is obviously bounded in some interval $x < t \leq \delta$—it tends to the limit $f'^+(x)$ as $t \to x^+$—and therefore the integral $\int_{x}^{x+\delta}\left|\frac{f(t)-f(x^+)}{t-x}\right|dt < +\infty$; similarly on the left. In fact, all one would need would be that in some interval $x < t \leq x+\delta$ the function satisfied an inequality of the form $|f(t)-f(x^+)| \leq (\text{const.})\cdot|t-x|^\alpha$, where $0 < \alpha \leq 1$, because then one would have[9]

$$\left|\frac{f(t)-f(x^+)}{t-x}\right| \leq \frac{(\text{const.})}{(t-x)^{1-\alpha}}$$

(9) Such a function is said to satisfy a *Hölder condition with exponent* α on the right (or left). Functions that may not be differentiable but do satisfy Hölder conditions occur frequently in classical potential theory. The (two-sided) Hölder conditions are "stronger than continuity but weaker than differentiability."

and clearly the integral over $(x, x+\delta]$ of the r. h. s. is finite: up to translation and a constant, it is the same as $\displaystyle\int_0^\delta \frac{dt}{t^{1-\alpha}} < \infty$. Thus functions like $f(x) = \begin{cases} \sqrt{x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$, even though they are not differentiable at $x = 0$, satisfy the Dini condition there (and trivially this function satisfies it everywhere else on $(-\pi, \pi)$).

In particular, we see that if $f(x)$ is Riemann-integrable, has period 2π and is differentiable at every point of $[-\pi, \pi]$, then its Fourier series converges to its value at every point $x \in [-\pi, \pi]$. This gives us only *pointwise* convergence; however, Strauss shows on pp. 135–136 that if $f \in \mathcal{C}^1([-\pi, \pi])$ (has a *continuous* derivative, necessarily also periodic with period 2π), then in fact the Fourier series of $f(x)$ converges to $f(x)$ *uniformly* on $[-\pi, \pi]$. We would like to take a different approach and show that if there is a Riemann-integrable function $g(x)$ (defined on the real line, with period 2π) such that $f(x) \equiv (\text{const.}) + \displaystyle\int_0^x g(t)\, dt$, then the Fourier series of $f(x)$ converges to $f(x)$ uniformly on $[-\pi, \pi]$. This is a much weaker assumption, satisfied by many functions that are continuous but fail to have a derivative at a certain discrete set of points, and on the way to proving this stronger theorem we can prove a fact that at this stage of the development "ought to be true" but which we have not as yet established.

2. Norm-Convergence of "All" Fourier Series: By "all" Fourier series we have to understand the Fourier series of all Riemann-integrable functions on $[-\pi, \pi]$. Here we shall need some preliminary lemmas whose relevance may not be immediately obvious. The proof will proceed by stages: we shall show first that the Fourier series of step functions converge to them in the r. m. s. norm, and then move to general Riemann-integrable functions in the same way we did when we proved the Riemann-Lebesgue lemma.

2.1. A Particular Function: Consider the function $g(x)$ defined by[10]

$$g(x) = \frac{x^2}{\pi^2} - \frac{x}{\pi} + \frac{1}{6}$$

on $[0, \pi]$ and extended to a function of period π "by making it repeat at intervals of length π." Because $g(0) = g(\pi)$, the resulting (extended) function is continuous, with left- and right-hand derivatives at every point (indeed, these differ only when x is a multiple of π). It follows from the results of the preceding section that the Fourier series of $g(x)$ converges to $g(x)$ at every $x \in \mathbb{R}$. Since $g(x)$ is an even function, the sine terms in its Fourier (trig) series have zero coefficients, and its cosine coefficients are given[11] by

$$\begin{aligned} A_0 &= \frac{2}{\pi}\int_0^\pi \left[\frac{x^2}{\pi^2} - \frac{x}{\pi} + \frac{1}{6}\right] dx = \frac{2}{3} - \frac{2}{2} + \frac{2}{6} = 0 \\ A_n &= \frac{2}{\pi}\int_0^\pi \left[\frac{x^2}{\pi^2} - \frac{x}{\pi} + \frac{1}{6}\right] \cos nx\, dx = \frac{2 \cdot [1 + \cos(n\pi)]}{n^2\pi^2} \end{aligned}$$

so (since the odd-indexed terms have zero coefficients, we may set $n = 2k$)

$$g(x) = \sum_{k=1}^\infty \frac{\cos 2kx}{k^2\pi^2}\ .$$

This equality holds for every $x \in \mathbb{R}$, and in fact (by comparison with the series $\sum_k 1/k^2$) the series converges to $g(x)$ *uniformly on* $\mathbb{R}$, though we shall not need that fact. In particular, for $x = 0$ we get the familiar relation

$$\frac{\pi^2}{6} = \sum_{k=1}^\infty \frac{1}{k^2}\ .$$

(10) The reader may recognize this as the periodic version of the second Bernoulli polynomial P_2 with the argument x/π plugged in.

(11) This computation is due to Maple.

2.2. The Fourier Series of a Pulse: Let $0 < a < \pi$ and consider the "symmetric unit pulse" function defined for $x \in [-\pi, \pi]$ by

$$f(x) = \begin{cases} 1 & \text{for } |x| < a \\ 1/2 & \text{for } |x| = a \\ 0 & \text{for } a < |x| < \pi \end{cases}$$

and extended to be periodic with period 2π. Since it is an even function its Fourier (trig) series is a cosine series, with coefficients

$$A_n = \frac{2}{\pi}\int_0^a \cos nx\, dx = \begin{cases} \dfrac{2a}{\pi} & \text{for } n = 0; \\ \dfrac{2\sin na}{n\pi} & \text{for } n > 0. \end{cases}$$

This function satisfies left- and right-hand Dini conditions at every x, and so we have

$$g(x) = \frac{a}{\pi} + \sum_{n=1}^{\infty} \frac{2\sin na}{n\pi} \cos nx \quad \text{(pointwise convergence)}$$

for every $x \in \mathbb{R}$; however, we would like to show that $g(x)$ is the r. m. s. limit of its Fourier series. The Parseval relation tells us that for that to be true it is necessary and sufficient that the relation

$$\int_{-\pi}^{\pi} g(x)^2\, dx = \left(\frac{A_0}{2}\right)^2 \cdot 2\pi + \sum_{k=1}^{\infty} A_k^2 \cdot \pi, \quad \text{or}$$

$$2a = \frac{a^2}{\pi^2} \cdot 2\pi + \sum_{k=1}^{\infty} \frac{4\sin^2 ka}{k^2\pi^2} \cdot \pi$$

should hold. Using the identity $\sin^2\theta = \dfrac{1-\cos 2\theta}{2}$, we see that this relation is logically equivalent to

$$2a = \frac{a^2}{\pi} \cdot 2 + \sum_{k=1}^{\infty} \frac{2(1-\cos 2ka)}{k^2\pi}$$

$$\frac{a^2}{\pi^2} - \frac{a}{\pi} + \frac{1}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2} = \sum_{k=1}^{\infty}\frac{\cos 2ka}{k^2\pi^2}$$

$$\frac{a^2}{\pi^2} - \frac{a}{\pi} + \frac{1}{6} = \sum_{k=1}^{\infty}\frac{\cos 2ka}{k^2\pi^2}\ ,$$

and since that relation (with x replaced by a) is exactly what we established in §**2.1** above, we have shown that **every symmetric pulse function is the r. m. s. limit of (the partial sums of) its Fourier series**.

2.3. From Pulses to Step Functions: This part of the argument we shall give in words, because writing it out symbolically would only obscure its simplicity. In view of the translation relations established in §**0.4** above and the geometrically obvious fact that every pulse (*i.e.*, function that takes a constant value c on some interval $[a, b] \subseteq [-\pi, \pi]$, is zero elsewhere in that interval, and is extended to have period 2π) is a scalar multiple of a translate of a symmetric unit pulse of the kind just considered in §**2.2** above, we see that every pulse is the r. m. s. limit of its Fourier series. Because changing the values of a function at a finite set of points of its domain does not change its integral, we may change the values of a pulse at the endpoints of the interval on which it is nonzero without changing its Fourier coefficients or the fact that it is the r. m. s. limit of its Fourier series. Now every step function(12) is a finite sum of pulses. It

(12) See the separate notes on the Riemann-Lebesgue lemma for a more complete discussion of step functions.

follows—since the Fourier coefficients of a finite sum of functions can be computed term-by-term[13] and then added—that **the Fourier series of any step function on $[-\pi, \pi]$ converges to that step function in the r. m. s. norm**.

2.4. From Step Functions to Riemann-Integrable Functions: We need one more little lemma and we can prove the main theorem of this section.

Lemma: Let $f(x)$ be a Riemann-integrable function on an interval $[a, b]$. Then for every $\eta > 0$ there exist a step function $s(x)$ for which $\|f - s\| = \sqrt{\int_a^b |f(x) - s(x)|^2\, dx} < \eta$.

Proof. It suffices to consider real-valued $f(x)$. Every (properly) Riemann-integrable function is bounded, so let $M \geq 0$ be such that $|f(x)| \leq M$ holds for all $x \in [a, b]$. It follows from the definition of the Riemann integral[14] that for every $\epsilon > 0$ there exists a step function $s(x)$ such that $s(x) \leq f(x)$ holds for $x \in [a, b]$ and such that $\int_a^b [f(x) - s(x)]\, dx < \epsilon$. We can take $s(x)$ such that $-M \leq s(x) \leq M$ also: $s(x) \leq f(x) \leq M$ would hold in all cases, and if on some interval $[x_{j-1}, x_j]$ we had taken $s(x) = c_j < -M$, we could replace c_j by $-M$ (because $-M \leq f(x)$ for all $x \in [a, b]$) while keeping $s(x) \leq f(x)$ and making $f(x) - s(x)$ (and thus its integral) smaller, if anything. But then $|f(x) - s(x)| \leq 2M$ on $[a, b]$, and consequently

$$\int_a^b |f(x) - s(x)|^2\, dx \leq 2M \cdot \int_a^b [f(x) - s(x)]\, dx\ . \tag{\$}$$

We know that given $\eta > 0$ we may find $s(x) \leq f(x)$ for which $\int_a^b [f(x) - s(x)]\, dx < \dfrac{\eta^2}{2M}$; the inequality ($) then tells us that the assertion of the lemma will hold for that $s(x)$.

Theorem: Let $f(x)$ be a function of period 2π that is Riemann-integrable on $[-\pi, \pi]$ and therefore on every finite interval. Then the (formal) Fourier series of $f(x)$ converges to $f(x)$ in the r. m. s. norm.

Proof. Let $\epsilon > 0$ be given; we shall show that there exists N_ϵ such that

$$N \geq N_\epsilon \Longrightarrow \left\| f(x) - \sum_{n=-N}^{N} \langle f, e^{inx} \rangle e^{inx} \right\| < \epsilon\ .$$

First, find a step function $s(x)$ for which $\|f - s\| < \dfrac{\epsilon}{2}$; this is possible by the immediately-preceding lemma. Let $c_n = \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} s(t) e^{-int}\, dt = \langle s, e^{inx} \rangle$. Because the Fourier series of step functions are known to converge in the r. m. s. norm to their corresponding functions, there exists N_ϵ such that

$$N \geq N_\epsilon \Longrightarrow \left\| s(x) - \sum_{n=-N}^{N} c_n e^{inx} \right\| < \frac{\epsilon}{2}\ .$$

By the triangle inequality for the r. m. s. norm, it follows that

$$N \geq N_\epsilon \Longrightarrow \left\| f(x) - \sum_{n=-N}^{N} c_n e^{inx} \right\| \leq \|f - s\| + \left\| s(x) - \sum_{n=-N}^{N} c_n e^{inx} \right\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\ .$$

---

(13) This is to say, the passages from functions to Fourier coefficients and from Fourier coefficients to partial sums of Fourier series are *linear* operations.

(14) See p. 2 of the notes on the Riemann-Lebesgue lemma for a discussion of this assertion.

The Bessel-Parseval relations imply (this was observation (1) of §**0.2** above) that the sum of the form $\sum_{n=-N}^{N} c_n e^{inx}$ that is *closest* to $f(x)$ in the r. m. s. norm is $\sum_{n=-N}^{N} \langle f, e^{inx}\rangle e^{inx}$. It follows that

$$N \geq N_\epsilon \Longrightarrow \left\| f(x) - \sum_{n=-N}^{N} \langle f, e^{inx}\rangle e^{inx} \right\| \leq \|f - s\| + \left\| s(x) - \sum_{n=-N}^{N} c_n e^{inx} \right\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\ ,$$

concluding the proof of the theorem.

3. Integration, Differentiation and Fourier Series: My feeling (and that of many people) is that looking at *integration* rather than *differentiation* is the way to gain insight into the relation between these operations of calculus and the convergence of Fourier series.

3.1. Fourier Coefficients of a Function vs. its Indefinite Integral: This is one of the places in which trig series make more sense than complex-exponential series and in which the interval $[0, 2\pi]$ is to be preferred to $[-\pi, \pi]$. Suppose $f(x)$ is a function defined for $x \in \mathbb{R}$ that is periodic with period 2π and Riemann-integrable on $[0, 2\pi]$ (and therefore on every finite interval). Let a_n and b_n denote its Fourier cosine and sine coefficients,

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos nt\,dt \quad \text{and} \quad b_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin nt\,dt$$

and suppose, for the moment, that $a_0 = 0$ {one can always replace $f(x)$ by $f(x) - \frac{a_0}{2}$, work with that new function, and then "put a_0 back" later, so this is not a very restrictive assumption}. Let $F(x)$ denote its "indefinite integral," or really its definite integral from 0 to x:

$$F(x) = \int_0^x f(t)\,dt\ .$$

Then the assumption $a_0 = 0$ implies that $F(2\pi) = 0 = F(0)$, so $F(x)$ is also periodic with period 2π (verification easy). We can compute the Fourier coefficients of $F(x)$ in terms of those of $f(x)$ by "reversing the order of integration," which is a valid operation[15] for Riemann-integrable functions. Letting A_n and B_n denote the Fourier cosine and sine coefficients of $F(x)$, we have

$$\begin{aligned} A_n &= \frac{1}{\pi}\int_0^{2\pi} F(x)\cos nx\,dx = \frac{1}{\pi}\int_0^{2\pi}\int_0^x f(t)\,dt\ \cos nx\,dx \\ &= \frac{1}{\pi}\int_0^{2\pi}\int_t^{2\pi} f(t)\ \cos nx\,dx\,dt = \frac{1}{\pi}\int_0^{2\pi} f(t)\,\frac{\sin 2n\pi - \sin nt}{n}\,dt \\ &= \frac{-1}{n}\cdot\frac{1}{\pi}\int_0^{2\pi} f(t)\sin nt\,dt = \frac{-1}{n}\,b_n \quad \text{for } n > 0. \end{aligned}$$

(Of course $A_0 = \frac{1}{\pi}\int_0^{2\pi} F(x)\,dx$ is not necessarily zero.) The computation for the sine coefficients is only slightly less unsubtle that that for the cosines:

$$\begin{aligned} B_n &= \frac{1}{\pi}\int_0^{2\pi} F(x)\sin nx\,dx = \frac{1}{\pi}\int_0^{2\pi}\int_0^x f(t)\,dt\ \sin nx\,dx \\ &= \frac{1}{\pi}\int_0^{2\pi}\int_t^{2\pi} f(t)\ \sin nx\,dx\,dt = \frac{1}{\pi}\int_0^{2\pi} f(t)\,\frac{\cos nt - \cos 2n\pi}{n}\,dt \\ &= \frac{1}{n}\cdot\left\{\frac{1}{\pi}\int_0^{2\pi} f(t)[\cos nt - 1]\,dt\right\} = \frac{1}{n}\,[a_n - a_0] = \frac{1}{n}\,a_n\ . \end{aligned}$$

(15) We omit the proof, but it's easy: as in (almost) all Riemann-integral situations, one approximates by step functions.

The effect of these computations is most easily understood by comparing the complete n-th harmonic term of the Fourier expansion of $f(x)$ with that of $F(x)$, expressing the A_n's and B_n's in terms of the a_n's and b_n's:

$$a_n \cos nx + b_n \sin nx \quad \text{vs.} \quad -\frac{a_n}{n} \sin nx + \frac{b_n}{n} \cos nx \ .$$

The complete n-th harmonic term ($n > 0$) of the Fourier expansion of $F(x)$ is obtained from that of $f(x)$ by formally integrating the trig functions, omitting the "$\cos n0$" terms that one might expect from the definite integral of $\sin nt$ from 0 to x.

The next step is to observe that **the Fourier series of $F(x)$ converges absolutely and uniformly on the interval** $[0, 2\pi]$ (and therefore, by the periodicity of everything in sight, **absolutely and uniformly on the entire real line**. The reason is that we can estimate the terms of this Fourier series is a very simple way: we know that $\displaystyle\sum_{n=1}^{\infty}(a_n^2 + b_n^2) \leq \frac{1}{\pi}\|f\|^2$. The Schwarz inequality—even for vectors in $\mathbb{R}^N$ or $\mathbb{C}^N$—then tells us that for every N

$$\sum_{n=1}^{N} \frac{1}{n}|a_n| \leq \sqrt{\sum_{n=1}^{N} a_n^2} \cdot \sqrt{\sum_{n=1}^{N} \frac{1}{n^2}} \leq \frac{1}{\sqrt{\pi}}\|f\| \cdot \frac{\pi}{\sqrt{6}} = \sqrt{\frac{\pi}{6}}\|f\|$$

and similarly for $\displaystyle\sum_{n=1}^{N} \frac{1}{n}|b_n|$. So the sums $\displaystyle\sum_{n=1}^{\infty} \frac{a_n}{n} \sin nx$ and $\displaystyle\sum_{n=1}^{\infty} \frac{b_n}{n} \cos nx$ converge absolutely, and they converge absolutely uniformly because the absolute values of their terms can be estimated by the corresponding terms of the constant series $\displaystyle\sum_{n=1}^{\infty} \frac{|a_n|}{n}$ and $\displaystyle\sum_{n=1}^{\infty} \frac{|b_n|}{n}$ respectively.

In view of the niceness of convergence of the Fourier series of $F(x)$ it would be dismaying if it converged to something other than $F(x)$, but of course that doesn't happen: applying the Schwarz inequality for integrals to the difference of two values of $F(x)$ (assume that $y < x$, since the situation is symmetric, and that the difference between them is $< 2\pi$, since we shall only be interested in the case in which they are close to each other), we get

$$|F(x) - F(y)| = \left|\int_y^x f(t)\,dt\right| \leq \sqrt{\int_y^x |f(t)|^2\,dt} \cdot \sqrt{\int_y^x 1\,dt} \leq \|f\| \cdot |x - y|^{1/2} \ .$$

This says that $F(x)$ satisfies a Hölder condition (with exponent $\alpha = 1/2$) at each point,(16) so by Dini's theorem its Fourier series converges to its value at every point $x \in [0, 2\pi]$ (and therefore at every point of the real line $\mathbb{R}$). We have proved:

Theorem: Let $f(x)$ be a function of period 2π, Riemann-integrable on $[0, 2\pi]$ with $\displaystyle\int_0^{2\pi} f(x)\,dx = 0$. Then term-by-term formal integration of its Fourier series yields the Fourier series of its indefinite integral: if $F(x) = \displaystyle\int_0^x f(t)\,dt$, then

$$f(x) = \lim_{N\to\infty} \sum_{n=1}^{N} [a_n \cos nx + b_n \sin nx] \quad \text{in } L^2\text{-norm yields}$$

$$F(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty}\left[\frac{a_n}{n} \sin nx - \frac{b_n}{n} \cos nx\right] \quad \text{with the series converging absolutely uniformly to } F(x).$$

(16) Actually $F(x)$ satisfies a Hölder condition with exponent $\alpha{=}1$ (also called a *Lipschitz condition*) at each point: if M is a bound for the function $f(x)$, so that $|f(x)| \leq M$ holds everywhere, then $|F(x) - F(y)| \leq M \cdot |x - y|$. The Hölder condition just proved, however, is the best one can do when one only knows that the improper integral of $|f(x)|^2$ is finite, a situation to be discussed briefly below.

If the mean value $\dfrac{a_0}{2} = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} f(x)\,dx$ of $f(x)$ is not zero, then by subtracting it off, applying the theorem just stated to $f(x) - \dfrac{a_0}{2}$, and then putting back the integrated term, one finds that

$$F(x) = \frac{A_0}{2} + \frac{a_0}{2}\,x + \sum_{n=1}^{\infty}\left[\frac{a_n}{n}\sin nx - \frac{b_n}{n}\cos nx\right] \quad \text{with the series converging absolutely uniformly.}$$

(It is trivial to fill in the details.) Note that $F(x)$ will be periodic if and only if $a_0 = 0$.

3.2. A Useful Example: Consider the function defined on $[0,1]$ by

$$P_1(x) = \begin{cases} x - \dfrac{1}{2} & \text{for } 0 < x < 1 \\ 0 & \text{for } x = 0 \text{ and } x = 1 \end{cases}$$

and extended to be periodic with period 1. It is easy to see (draw a picture! although this fact can also be verified by pure computation) that this is an odd function, so its Fourier trig series has only sine terms. Its coefficients are given by

$$b_n = 2\int_0^1 \left(x - \frac{1}{2}\right)\sin 2\pi nx\,dx = \frac{-1}{n\pi}\ ,$$

and since it is evidently Riemann-integrable and satisfies the Dini condition from the right and left at each $x \in \mathbb{R}$, we have

$$P_1(x) = \lim_{N\to\infty}\sum_{n=1}^{N}\frac{-1}{n\pi}\,\sin 2\pi nx$$

pointwise and in L^2-norm, but not uniformly (a uniformly convergent series of continuous functions would have to converge to a continuous function, and $P_1(x)$ is discontinuous at every integer value of x). The indefinite integral of $P_1(x)$ is $\dfrac{x^2}{2} - \dfrac{x}{2}$ and the mean of this function is $\dfrac{1}{6} - \dfrac{1}{4} = \dfrac{-1}{12}$, so we have

$$\frac{x^2}{2} - \frac{x}{2} = \frac{-1}{12} + \sum_{n=1}^{\infty}\frac{1}{2n^2\pi^2}\,\cos 2\pi nx$$

uniformly on $[0,1]$, with the r. h. side converging absolutely uniformly on $\mathbb{R}$ to the periodic extension with period 1 of the l. h. side. (Of course we met this function—scaled to have period π—in §§**2.1** and **2.2** above.) It is a nice exercise for the reader to see how this integration can be repeated indefinitely often. (It is possible to find a generating function for the results.)

Aside from the usefulness of knowing the sum of this series, this example shows that even a discontinuous derivative, or one that fails to exist at a discrete set of points, can have the "right" Fourier series. The key question is: is the derivative Riemann-integrable, and is its indefinite Riemann integral equal to the function one started with? In this case, we see that $P_2(x)$ fails to have a derivative when x is an integer, and the function defined at non-integer points by $P_2'(x)$ is discontinuous. But because $P_2(x)$ is the indefinite integral of its (Riemann-integrable, and discontinuous only at a discrete set of points) derivative, one can see (after the fact) that formal term-by-term differentiation of its Fourier series is justified. *This* is the kind of relation between the Fourier series of a function and that of its derivative that one wants to have: one does not want to restrict oneself to *continuously* differentiable functions.

4. Touching Base with Strauss: Having entertained myself by giving a self-contained treatment of some elementary convergence questions for Fourier series, I must now check that I have covered everything that Strauss did.[17] Let's pick our way through Chapter 5. In §**5.4**, Strauss doesn't claim to have proved

[17] Fighting the textbook is bad enough, but not doing at least as much as it does is much worse.

Theorem 1 and neither do I, except in the Dirichlet- or Neumann-boundary-condition cases. He doesn't claim to have proved **Theorem 2** and neither do I, although I have proved his side remark about **Theorem 2**, namely: If $f(x)$ and $f'(x)$ are continuous and periodic (with some period 2ℓ), then the (full) Fourier series of $f(x)$ converges absolutely uniformly. I can say I have proved this because if $f'(x)$ is continuous then $f(x) = (\text{const.}) + \int_0^x f'(t)\,dt$ (the fundamental theorem of calculus). By proving Dini's theorem, I have justified Strauss's assertions **Theorem 4** and **Theorem 4$^\infty$**. I have also verified (everybody gives about the same proofs of these things) the assertions made under the heading **The $\mathbf{L^2}$ Theory**, and the same goes for the assertions of §**5.5** as far as but not including the material headed **The Gibbs Phenomenon**.

4.1. The Hypothesis $\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$: This treatment has concentrated on "properly" Riemann-integrable functions. However, many of the resuts extend to functions whose squares are "improperly" Riemann-integrable, for example, the function $f(x)$ defined by $|x|^{-1/4}$ on $[-\pi, \pi]$ and extended to be periodic with period 2π on $\mathbb{R}$. Such functions have well-defined Fourier coefficients, and their Fourier series converge to the function in the r. m. s. norm: one approximates the given function in the r. m. s. norm by a properly Riemann-integrable function (the ability to do this is built into the definition of improper Riemann integrals) and then one can just use the same argument with step functions that we gave above. At points where such functions satisfy the Dini condition, their Fourier series converge to the average of the one-sided limits (in this case of this example, everywhere but at $x = 2k\pi$ where $k \in \mathbb{Z}$). But one would be hard put to compute the Fourier (cosine) coefficients of this $f(x)$; its Fourier series is going to converge so slowly that it would be nearly useless for approximation purposes; and if one is going to work with these "square-summable functions," there is really nothing to be gained by sticking with Riemann's definition of the integral. One might as well use the Lebesgue theory of integration, which gives all of the results obtainable with the Riemann integral, gives the same values to the integrals of the elementary functions, and simplifies a lot of the proofs (of course, there is a conservation-of-work principle involved: it's a lot of work to set up Lebesgue integration theory!).

My purpose in writing up these notes has been to give a reasonably rigorous and self-contained treatment of convergence of the Fourier series of "garden-variety" functions, those that are Riemann integrable and have left- and right-hand limits and derivatives at each point (for pointwise convergence) or the indefinite integrals of Riemann-integrable functions (for uniform convergence). I claim that it takes very little more work than what Strauss's book does, gets somewhat more inclusive results, and requires little to be taken on faith. I don't make any claims of originality.