Academic Press

Encyclopedia of Physical Science and Technology

Fourier Series

James S. Walker
Department of Mathematics
University of Wisconsin–Eau Claire
Eau Claire, WI 54702–4004

Phone: 715–836–3301
Fax: 715–836–2924
e-mail: walkerjs@uwec.edu
I. Introduction
Fourier series has long provided one of the principal methods of analysis for mathematical physics, engineering, and signal processing. It has spurred generalizations...
and applications that continue to develop right up to the present. While the original theory of Fourier series applies to periodic functions occurring in wave motion, such as with light and sound, its generalizations often relate to wider settings, such as the time-frequency analysis underlying the recent theories of wavelet analysis and local trigonometric analysis.

II. Historical background

There are antecedents to the notion of Fourier series in the work of Euler and D. Bernoulli on vibrating strings, but the theory of Fourier series truly began with the profound work of Fourier on heat conduction at the beginning of the 19th century. In [5], Fourier deals with the problem of describing the evolution of the temperature \( T(x, t) \) of a thin wire of length \( \pi \), stretched between \( x = 0 \) and \( x = \pi \), with a constant zero temperature at the ends: \( T(0, t) = 0 \) and \( T(\pi, t) = 0 \). He proposed that the initial temperature \( T(x, 0) = f(x) \) could be expanded in a series of sine functions:

\[
f(x) = \sum_{n=1}^{\infty} b_n \sin nx
\]

with

\[
b_n = \frac{2}{\pi} \int_{0}^{\pi} f(x) \sin nx \, dx.
\]

Although Fourier did not give a convincing proof of convergence of the infinite series in Eq. (1), he did offer the conjecture that convergence holds for an “arbitrary” function \( f \). Subsequent work by Dirichlet, Riemann, Lebesgue, and others, throughout the next two hundred years was needed to delineate precisely which functions were expandable in such trigonometric series. Part of this work entailed giving a precise definition of function (Dirichlet), and showing that the integrals in Eq. (2) are properly defined (Riemann and Lebesgue). Throughout this article we shall state results that are always true when Riemann integrals are used (except for Sec. V where we need to use results from the theory of Lebesgue integrals).

In addition to positing (1) and (2), Fourier argued that the temperature \( T(x, t) \) is a solution to the following heat equation with boundary conditions:

\[
\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}, \quad 0 < x < \pi, \ t > 0
\]

\[
T(0, t) = T(\pi, t) = 0, \quad t \geq 0
\]

\[
T(x, 0) = f(x), \quad 0 \leq x \leq \pi.
\]

Making use of (1), Fourier showed that the solution \( T(x, t) \) satisfies

\[
T(x, t) = \sum_{n=1}^{\infty} b_n e^{-n^2 t} \sin nx.
\]
This was the first example of the use of Fourier series to solve boundary value problems in partial differential equations. To obtain (3), Fourier made use of D. Bernoulli’s method of separation of variables, which is now a standard technique for solving boundary value problems.

A good, short introduction to the history of Fourier series can be found in [4]. Besides his many mathematical contributions, Fourier has left us with one of the truly great philosophical principles: “The deep study of nature is the most fruitful source of knowledge.”

### III. Definition of Fourier series

The Fourier sine series, defined in Eq.s (1) and (2), is a special case of a more general concept: the Fourier series for a periodic function. Periodic functions arise in the study of wave motion, when a basic waveform repeats itself periodically. Such periodic waveforms occur in musical tones, in the plane waves of electromagnetic vibrations, and in the vibration of strings. These are just a few examples. Periodic effects also arise in the motion of the planets, in ac-electricity, and (to a degree) in animal heartbeats.

A function \( f \) is said to have period \( P \) if \( f(x + P) = f(x) \) for all \( x \). For notational simplicity, we shall restrict our discussion to functions of period \( 2\pi \). There is no loss of generality in doing so, since we can always use a simple change of scale \( x = (P/2\pi)t \) to convert a function of period \( P \) into one of period \( 2\pi \).

If the function \( f \) has period \( 2\pi \), then its Fourier series is

\[
c_0 + \sum_{n=1}^{\infty} \left\{ a_n \cos nx + b_n \sin nx \right\}
\]

with Fourier coefficients \( c_0, a_n, \) and \( b_n \) defined by the integrals

\[
c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx \tag{5}
\]

\[
a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \tag{6}
\]

\[
b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \tag{7}
\]

[Note: The sine series defined by (1) and (2) is a special instance of Fourier series. If \( f \) is initially defined over the interval \([0, \pi]\), then it can be extended to \([-\pi, \pi]\) (as an odd function) by letting \( f(-x) = -f(x) \), and then extended periodically with period \( P = 2\pi \). The Fourier series for this odd, periodic function reduces to the sine series in Eq.s (1) and (2), because \( c_0 = 0 \), each \( a_n = 0 \), and each \( b_n \) in Eq. (7) is equal to the \( b_n \) in Eq. (2).]
Fourier series

It is more common nowadays to express Fourier series in an algebraically simpler form involving complex exponentials. Following Euler, we use the fact that the complex exponential $e^{i\theta}$ satisfies $e^{i\theta} = \cos \theta + i \sin \theta$. Hence

$$\cos \theta = \frac{1}{2} (e^{i\theta} + e^{-i\theta}),$$

$$\sin \theta = \frac{1}{2i} (e^{i\theta} - e^{-i\theta}).$$

From these equations, it follows by elementary algebra that Formulas (5)–(7) can be rewritten (by rewriting each term separately) as

$$c_0 + \sum_{n=1}^{\infty} \left\{ c_n e^{in\pi/2} + c_n e^{-in\pi/2} \right\}$$

(8)

with $c_n$ defined for all integers $n$ by

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-in\pi/2} dx.$$  

(9)

The series in (8) is usually written in the form

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}.$$  

(10)

We now consider a couple of examples. First, let $f_1$ be defined over $[-\pi, \pi]$ by

$$f_1(x) = \begin{cases} 1 & \text{if } |x| < \pi/2 \\ 0 & \text{if } \pi/2 \leq |x| \leq \pi \end{cases}$$

and have period $2\pi$. The graph of $f_1$ is shown in Fig. 1; it is called a square wave in electric circuit theory. The constant $c_0$ is

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_1(x) \, dx = \frac{1}{2 \pi} \int_{-\pi/2}^{\pi/2} 1 \, dx = \frac{1}{2}.$$  

While, for $n \neq 0$,

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_1(x) e^{-in\pi/2} \, dx = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} e^{-in\pi/2} \, dx$$

$$= \frac{1}{2\pi} \left[ -\frac{e^{-in\pi/2} - e^{in\pi/2}}{in\pi} \right] = \frac{\sin(n\pi/2)}{n\pi}.$$
Thus, the Fourier series for this square wave is

\[
\frac{1}{2} + \sum_{n=1}^{\infty} \frac{\sin(n\pi/2)}{n\pi} \left( e^{inx} + e^{-inx} \right) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{2\sin(n\pi/2)}{n\pi} \cos nx, \tag{11}
\]

Second, let \( f_2(x) = x^2 \) over \([-\pi, \pi]\) and have period \(2\pi\). See Fig. 2. We shall refer to this wave as a parabolic wave. This parabolic wave has \( c_0 = \pi^2 / 3 \) and \( c_n \), for \( n \neq 0 \), is

\[
c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 e^{-inx} \, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 \cos nx \, dx - \frac{i}{2\pi} \int_{-\pi}^{\pi} x^2 \sin nx \, dx = \frac{2(-1)^n}{n^2} \]

after an integration by parts. The Fourier series for this function is then

\[
\frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{2(-1)^n}{n^2} \left( e^{inx} + e^{-inx} \right)
\]
We will discuss the convergence of these Fourier series, to \( f_1 \) and \( f_2 \) respectively, in Section IV.

Returning to the general Fourier series in Eq. (10), we shall now discuss some ways of interpreting this series. A complex exponential \( e^{inx} = \cos(nx) + i \sin(nx) \) has a smallest period of \( \frac{2\pi}{n} \). Consequently it is said to have a frequency of \( \frac{n}{2\pi} \), because the form of its graph over the interval \([0, \frac{2\pi}{n}]\) is repeated \( n/2\pi \) times within each unit-length. Therefore, the integral in (9) that defines the Fourier coefficient \( c_n \) can be interpreted as a correlation between \( f \) and a complex exponential with a precisely located frequency of \( n/2\pi \). Thus the whole collection of these integrals, for all integers \( n \), specifies the frequency content of \( f \) over the set of frequencies \( \{n/2\pi\}_{n=-\infty}^{\infty} \). If the series in (10) converges to \( f \), i.e., if we can write

\[
 f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx},
\]

then \( f \) is being expressed as a superposition of elementary functions \( c_n e^{inx} \) having frequency \( n/2\pi \) and amplitude \( c_n \). [The validity of Eq. (13) will be discussed in the next section.] Furthermore, the correlations in Eq. (9) are independent of each
other in the sense that correlations between distinct exponentials are zero:

\[ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx} \, dx = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n. \end{cases} \]  

(14)

This equation is called the orthogonality property of complex exponentials.

The orthogonality property of complex exponentials can be used to give a derivation of Eq. (9). Multiplying Eq. (13) by \( e^{-imx} \) and integrating term-by-term from \(-\pi\) to \(\pi\), we obtain

\[ \int_{-\pi}^{\pi} f(x) e^{-imx} \, dx = \sum_{n=-\infty}^{\infty} c_n \int_{-\pi}^{\pi} e^{inx} e^{-imx} \, dx. \]

By the orthogonality property, this leads to

\[ \int_{-\pi}^{\pi} f(x) e^{-imx} \, dx = 2\pi c_m, \]

which justifies (in a formal, non-rigorous way) the definition of \( c_n \) in Eq. (9).

We close this section by discussing two important properties of Fourier coefficients, Bessel’s inequality and the Riemann-Lebesgue lemma.

**Theorem 1 (Bessel’s Inequality)** If \( \int_{-\pi}^{\pi} |f(x)|^2 \, dx \) is finite, then

\[ \sum_{n=-\infty}^{\infty} |c_n|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx. \]  

(15)

Bessel’s inequality can be proved easily. In fact, we have

\[
0 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx - \sum_{n=-N}^{N} |c_n e^{inx}|^2 \, dx \\
= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x) - \sum_{m=-N}^{N} c_m e^{imx})(f(x) - \sum_{n=-N}^{N} c_n e^{-inx}) \, dx.
\]

Multiplying out the last integrand above, and making use of Eq.s (9) and (14), we obtain

\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx - \sum_{n=-N}^{N} |c_n|^2 \\
= \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx - \sum_{n=-N}^{N} |c_n|^2.
\]  

(16)
Fourier series

Thus, for all $N$

$$\sum_{n=-N}^{N} |c_n|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx \tag{17}$$

and Bessel’s inequality (15) follows by letting $N \to \infty$.

Bessel’s inequality has a physical interpretation. If $f$ has finite energy, in the sense that the right side of (15) is finite, then the sum of the moduli-squared of the Fourier coefficients is also finite. In Sec. V, we shall see that the inequality in Eq. (15) is actually an equality, which says that the sum of the moduli-squared of the Fourier coefficients is precisely the same as the energy of $f$.

Because of Bessel’s inequality, it follows that

$$\lim_{|n| \to \infty} c_n = 0 \tag{18}$$

holds whenever $\int_{-\pi}^{\pi} |f(x)|^2 \, dx$ is finite. The Riemann-Lebesgue lemma says that Eq. (18) holds in the following more general case:

**Theorem 2 (Riemann-Lebesgue Lemma)** If $\int_{-\pi}^{\pi} |f(x)| \, dx$ is finite, then Eq. (18) holds.

One of the most important uses of the Riemann-Lebesgue lemma is in proofs of some basic pointwise convergence theorems, such as the ones described in the next section.

For further discussions of the definition of Fourier series, Bessel’s inequality, and the Riemann-Lebesgue lemma, see [7] or [12]

**IV. Convergence of Fourier series**

There are many ways to interpret the meaning of Eq. (13). Investigations into the types of functions allowed on the left side of (13), and the kinds of convergence considered for its right side, have fueled mathematical investigations by such luminaries as Dirichlet, Riemann, Weierstrass, Lipschitz, Lebesgue, Fejér, Gelfand, and Schwartz. In short, convergence questions for Fourier series have helped lay the foundations and much of the superstructure of mathematical analysis.

The three types of convergence that we shall describe here are pointwise, uniform, and norm convergence. We shall discuss the first two types in this section, and take up the third type in the next section.

All convergence theorems are concerned with how the partial sums

$$S_N(x) := \sum_{n=-N}^{N} c_n e^{inx}$$
converge to \( f(x) \). That is, does \( \lim_{N \to \infty} S_N = f(x) \) hold in some sense?

The question of pointwise convergence, for example, concerns whether

\[
\lim_{N \to \infty} S_N(x_0) = f(x_0)
\]

holds for each fixed \( x \)-value \( x_0 \). If \( \lim_{N \to \infty} S_N(x_0) \) does equal \( f(x_0) \), then we say that the Fourier series for \( f \) converges to \( f(x_0) \) at \( x_0 \).

We shall now state the simplest pointwise convergence theorem for which an elementary proof can be given. This theorem assumes that a function is Lipschitz at each point where convergence occurs. A function is said to be Lipschitz at a point \( x_0 \) if, for some positive constant \( A \),

\[
|f(x) - f(x_0)| \leq A |x - x_0|
\]

(19)
holds for all \( x \) near \( x_0 \) (i.e., \( |x - x_0| < \delta \) for some \( \delta > 0 \)). It is easy to see, for instance, that the square wave function \( f_1 \) is Lipschitz at all of its continuity points.

The inequality in (19) has a simple geometric interpretation. Since both sides are 0 when \( x = x_0 \), this inequality is equivalent to

\[
\left|\frac{f(x) - f(x_0)}{x - x_0}\right| \leq A
\]

(20)
for all \( x \) near \( x_0 \) (and \( x \neq x_0 \)). Inequality (20) simply says that the difference quotients of \( f \) (i.e., the slopes of its secants) near \( x_0 \) are bounded. With this interpretation, it is easy to see that the parabolic wave \( f_2 \) is Lipschitz at all points. More generally, if \( f \) has a derivative at \( x_0 \) (or even just left-hand and right-hand derivatives), then \( f \) is Lipschitz at \( x_0 \).

We can now state and prove a simple convergence theorem.

**Theorem 3** Suppose \( f \) has period \( 2\pi \), that \( \int_{-\pi}^{\pi} |f(x)| \, dx \) is finite, and that \( f \) is Lipschitz at \( x_0 \). Then the Fourier series for \( f \) converges to \( f(x_0) \) at \( x_0 \).

To prove this theorem, we assume that \( f(x_0) = 0 \). There is no loss of generality in doing so, since we can always subtract the constant \( f(x_0) \) from \( f(x) \). Define the function \( g \) by \( g(x) = f(x)/(e^{ix} - e^{ix_0}) \). This function \( g \) has period \( 2\pi \). Furthermore, \( \int_{-\pi}^{\pi} |g(x)| \, dx \) is finite, because the quotient \( f(x)/(e^{ix} - e^{ix_0}) \) is bounded in magnitude for \( x \) near \( x_0 \). In fact, for such \( x \),

\[
\left|\frac{f(x)}{e^{ix} - e^{ix_0}}\right| = \left|\frac{f(x) - f(x_0)}{e^{ix} - e^{ix_0}}\right| \leq A \left|\frac{x - x_0}{e^{ix} - e^{ix_0}}\right|
\]
Fourier series

and \((x - x_0)/(e^{ix} - e^{ix_0})\) is bounded in magnitude, because it tends to the reciprocal of the derivative of \(e^{ix}\) at \(x_0\).

If we let \(d_n\) denote the \(n\)th Fourier coefficient for \(g(x)\), then we have \(c_n = d_{n-1} - d_n e^{ix_0}\) because \(f(x) = g(x)(e^{ix} - e^{ix_0})\). The partial sum \(S_N(x_0)\) then telescopes:

\[
S_N(x_0) = \sum_{n=-N}^{N} c_n e^{inx_0} = d_{-N-1} e^{-iN x_0} - d_N e^{i(N+1)x_0}.
\]

Since \(d_n \to 0\) as \(|n| \to \infty\), by the Riemann-Lebesgue lemma, we conclude that \(S_N(x_0) \to 0\). This completes the proof.

It should be noted that for the square wave \(f_1\) and the parabolic wave \(f_2\), it is not necessary to use the general Riemann-Lebesgue lemma stated above. That is because for those functions it is easy to see that \(\int_{-\pi}^{\pi} |g(x)|^2 \, dx\) is finite for the function \(g\) defined in the proof of Theorem 3. Consequently, \(d_n \to 0\) as \(|n| \to \infty\) follows from Bessel’s inequality for \(g\).

In any case, Theorem 3 implies that the Fourier series for the square wave \(f_1\) converges to \(f_1\) at all of its points of continuity. It also implies that the Fourier series for the parabolic wave \(f_2\) converges to \(f_2\) at all points. While this may settle matters (more or less) in a pure mathematical sense for these two waves, it is still important to examine specific partial sums in order to learn more about the nature of their convergence to these waves.

For example, in Fig. 3 we show a graph of the partial sum \(S_{100}\) superimposed on the square wave. Although Theorem 3 guarantees that \(S_N \to f_1\) as \(N \to \infty\) at each continuity point, Fig. 3 indicates that this convergence is at a rather slow rate. The partial sum \(S_{100}\) differs significantly from \(f_1\). Near the square wave’s jump discontinuities, for example, there is a severe spiking behavior called Gibbs’ phenomenon (see Fig. 4). This spiking behavior does not go away as \(N \to \infty\), although the width of the spike does tend to zero. In fact, the peaks of the spikes overshoot the square wave’s value of 1, tending to a limit of about 1.09. The partial sum also oscillates quite noticeably about the constant value of the square wave at points away from the discontinuities. This is known as ringing.

These defects do have practical implications. For instance, oscilloscopes—which generate wave forms as combinations of sinusoidal waves over a limited range of frequencies—cannot use \(S_{100}\), or any partial sum \(S_N\), to produce a square wave. We shall see, however, in Sec. VI that a clever modification of a partial sum does produce an acceptable version of a square wave.

The cause of ringing and Gibbs’ phenomenon for the square wave is a rather slow convergence to zero of its Fourier coefficients (at a rate comparable to \(|n|^{-1}\)).
In the next section, we shall interpret this in terms of energy and show that a partial sum like $S_{100}$ does not capture a high enough percentage of the energy of the square wave $f_1$.

In contrast, the Fourier coefficients of the parabolic wave $f_2$ tend to zero more rapidly (at a rate comparable to $n^{-2}$). Because of this, the partial sum $S_{100}$ for $f_2$ is a much better approximation to the parabolic wave (see Fig. 5). In fact, its partial sums $S_N$ exhibit the phenomenon of uniform convergence.

We say that the Fourier series for a function $f$ converges uniformly to $f$ if

$$\lim_{N \to \infty} \left\{ \max_{x \in [-\pi, \pi]} |f(x) - S_N(x)| \right\} = 0. \quad (21)$$

This equation says that, for large enough $N$, we can have the maximum distance between the graphs of $f$ and $S_N$ as small as we wish. Fig. 5 is a good illustration of this for the parabolic wave.

We can verify Eq. (21) for the parabolic wave as follows. By Eq. (12) we have

$$|f_2(x) - S_N(x)| = \left| \sum_{n=N+1}^{\infty} \frac{4(-1)^n}{n^2} \cos nx \right|$$

$$\leq \sum_{n=N+1}^{\infty} \left| \frac{4(-1)^n}{n^2} \cos nx \right|$$
Fourier series

Figure 4: Gibbs’ phenomenon and ringing for square wave.

\[ \leq \sum_{n=N+1}^{\infty} \frac{4}{n^2}. \]

Consequently

\[ \max_{x \in [-\pi, \pi]} |f_2(x) - S_N(x)| \leq \sum_{n=N+1}^{\infty} \frac{4}{n^2} \]

\[ \to 0 \quad \text{as} \quad N \to \infty \]

and thus Eq. (21) holds for the parabolic wave \( f_2 \).

Uniform convergence for the parabolic wave is a special case of a more general theorem. We shall say that \( f \) is uniformly Lipschitz if Eq. (19) holds for all points using the same constant \( A \). For instance, it is not hard to show that a continuously differentiable, periodic function is uniformly Lipschitz.

**Theorem 4** Suppose that \( f \) has period \( 2\pi \) and is uniformly Lipschitz at all points, then the Fourier series for \( f \) converges uniformly to \( f \).

A remarkably simple proof of this theorem is described in [6]. More general uniform convergence theorems are discussed in [15].

Theorem 4 applies to the parabolic wave \( f_2 \), but it does not apply to the square wave \( f_1 \). In fact, the Fourier series for \( f_1 \) cannot converge uniformly to \( f_1 \). That
is because a famous theorem of Weierstrass says that a uniform limit of continuous functions (like the partial sums \( S_N \)) must be a continuous function (which \( f_1 \) is certainly not). The Gibbs’ phenomenon for the square wave is a conspicuous failure of uniform convergence for its Fourier series.

Gibbs’ phenomenon and ringing, as well as many other aspects of Fourier series, can be understood via an integral form for partial sums discovered by Dirichlet. This integral form is

\[
S_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - t) D_N(t) \, dt
\]  

(22)

with kernel \( D_N \) defined by

\[
D_N(t) = \frac{\sin (N + 1/2)t}{\sin(t/2)}.
\]

(23)

This formula is proved in almost all books on Fourier series (see, for instance, [7], [12], or [16]). The kernel \( D_N \) is called Dirichlet’s kernel. In Fig. 6 we have graphed \( D_{20} \).

The most important property of Dirichlet’s kernel is that, for all \( N \),

\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) \, dt = 1.
\]
From Eq. (23) we can see that the value of follow from cancellation of signed areas, and also that the contribution of the main lobe centered at (see Fig. 6) is significantly greater than 1 (about 1.09 in value).

From the facts just cited, we can explain the origin of ringing and Gibbs’ phenomenon for the square wave. For the square wave function \( f_1 \), Eq. (22) becomes

\[
S_N(x) = \frac{1}{2\pi} \int_{x-\pi/2}^{x+\pi/2} D_N(t) \, dt. \tag{24}
\]

As \( x \) ranges from \(-\pi\) to \( \pi \), this formula shows that \( S_N(x) \) is proportional to the signed area of \( D_N \) over an interval of length \( \pi \) centered at \( x \). By examining Fig. 6, which is a typical graph for \( D_N \), it is then easy to see why there is ringing in the partial sums \( S_N \) for the square wave. Gibbs’ phenomenon is a bit more subtle, but also results from Eq. (24). When \( x \) nears a jump discontinuity, the central lobe of \( D_N \) is the dominant contributor to the integral in Eq. (24), resulting in a spike which overshoots the value of 1 for \( f_1 \) by about 9%.

Our final pointwise convergence theorem was, in essence, the first to be proved. It was established by Dirichlet using the integral form for partial sums in Eq. (22). We shall state this theorem in a stronger form first proved by Jordan.

**Theorem 5** If \( f \) has period \( 2\pi \) and has bounded variation on \([0, 2\pi]\), then the
Fourier series for $f$ converges at all points. In fact, for all $x$-values,

$$\lim_{N \to \infty} S_N(x) = \frac{1}{2} [f(x+) + f(x-)].$$

This theorem is too difficult to prove in the limited space we have here (see [16]). A simple consequence of Theorem 5 is that the Fourier series for the square wave $f_1$ converges at its discontinuity points to $1/2$ [although this can also be shown directly by substitution of $x = \pm \pi/2$ into the series in (11)].

We close by mentioning that the conditions for convergence, such as Lipschitz or bounded variation, cited in the theorems above cannot be dispensed with entirely. For instance, Kolmogorov gave an example of a period $2\pi$ function (for which $\int_{-\pi}^{\pi} |f(x)| \, dx$ is finite) that has a Fourier series which fails to converge at every point.

More discussion of pointwise convergence can be found in [12], [15], or [16].

V. Convergence in norm

Perhaps the most satisfactory notion of convergence for Fourier series is convergence in $L^2$-norm (also called $L^2$-convergence), which we shall define in this section. One of the great triumphs of the Lebesgue theory of integration is that it yields necessary and sufficient conditions for $L^2$-convergence. There is also an interpretation of $L^2$-norm in terms of a generalized Euclidean distance and this gives a satisfying geometric flavor to $L^2$-convergence of Fourier series. By interpreting the square of $L^2$-norm as a type of energy, there is an equally satisfying physical interpretation of $L^2$-convergence. The theory of $L^2$-convergence has led to fruitful generalizations such as Hilbert space theory and norm convergence in a wide variety of function spaces.

To introduce the idea of $L^2$-convergence, we first examine a special case. By Theorem 4, the partial sums of a uniformly Lipschitz function $f$ converge uniformly to $f$. Since that means that the maximum distance between the graphs of $S_N$ and $f$ tends to 0 as $N \to \infty$, it follows that

$$\lim_{N \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - S_N(x)|^2 \, dx = 0. \quad (25)$$

This result motivates the definition of $L^2$-convergence.

If $g$ is a function for which $|g|^2$ has a finite Lebesgue integral over $[-\pi, \pi]$, then we say that $g$ is an $L^2$-function, and we define its $L^2$-norm $\|g\|_2$ by

$$\|g\|_2 = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(x)|^2 \, dx}.$$
We can then rephrase Eq. (25) as saying that $\|f - S_{N}\|_2 \to 0$ as $N \to \infty$. In other words, the Fourier series for $f$ converges to $f$ in $L^2$-norm. The following theorem generalizes this result to all $L^2$-functions (see [11] for a proof).

**Theorem 6** If $f$ is an $L^2$-function, then its Fourier series converges to $f$ in $L^2$-norm.

Theorem 6 says that Eq. (25) holds for every $L^2$-function $f$. Combining this with Eq. (16), we obtain Parseval’s equality:

$$\sum_{n=-\infty}^{\infty} |c_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx. \quad (26)$$

Parseval’s equation has a useful interpretation in terms of energy. It says that the energy of the set of Fourier coefficients, defined to be equal to the left side of Eq. (26), is equal to the energy of the function $f$, defined by the right side of Eq. (26).

The $L^2$-norm can be interpreted as a generalized Euclidean distance. To see this take square roots of both sides of Eq. (26): $\sqrt{\sum |c_n|^2} = \|f\|_2$. The left side of this equation is interpreted as a Euclidean distance in an (infinite-dimensional) coordinate space, hence the $L^2$-norm $\|f\|_2$ is equivalent to such a distance.

As examples of these ideas, let’s return to the square wave and parabolic wave. For the square wave $f_1$, we find that

$$\|f_1 - S_{100}\|_2^2 = \sum_{|n| > 100} \sin^2(n\pi/2) \left(\frac{(n\pi)^2}{n\pi}\right)^2$$

$$= 1.0 \times 10^{-3}.$$  

Likewise, for the parabolic wave $f_2$, we have $\|f_2 - S_{100}\|_2^2 = 2.6 \times 10^{-6}$. These facts show that the energy of the parabolic wave is almost entirely contained in the partial sum $S_{100}$; their energy difference is almost three orders of magnitude smaller than in the square wave case. In terms of generalized Euclidean distance, we have $\|f_2 - S_{100}\|_2 = 1.6 \times 10^{-3}$ and $\|f_1 - S_{100}\|_2 = 3.2 \times 10^{-2}$, showing that the partial sum is an order of magnitude closer for the parabolic wave.

Theorem 6 has a converse, known as the Riesz-Fischer theorem.

**Theorem 7 (Riesz-Fischer)** If $\sum |c_n|^2$ converges, then there exists an $L^2$-function $f$ having $\{c_n\}$ as its Fourier coefficients.

This theorem is proved in [11]. Theorem 6 and the Riesz-Fischer theorem combine to give necessary and sufficient conditions for $L^2$-convergence of Fourier series,
conditions which are remarkably easy to apply. This has made $L^2$-convergence into the most commonly used notion of convergence for Fourier series.

These ideas for $L^2$-norms partially generalize to the case of $L^p$-norms. Let $p$ be real number satisfying $p \geq 1$. If $g$ is a function for which $|g|^p$ has a finite Lebesgue integral over $[-\pi, \pi]$, then we say that $g$ is an $L^p$-function, and we define its $L^p$-norm $\|g\|_p$ by

$$\|g\|_p = \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(x)|^p \, dx \right]^{1/p}.$$  

If $\|f - S_N\|_p \to 0$, then we say that the Fourier series for $f$ converges to $f$ in $L^p$-norm. The following theorem generalizes Theorem 6 (see [7] for a proof).

**Theorem 8** If $f$ is an $L^p$-function for $p > 1$, then its Fourier series converges to $f$ in $L^p$-norm.

Notice that the case of $p = 1$ is not included in Theorem 8. The example of Kolmogorov cited at the end of Sec. IV shows that there exist $L^1$-functions whose Fourier series do not converge in $L^1$-norm. For $p \neq 2$, there are no simple analogs of either Parseval’s equality or the Riesz-Fischer theorem (which say that we can characterize $L^2$-functions by the magnitude of their Fourier coefficients). Some partial analogs of these latter results for $L^p$-functions, when $p \neq 2$, are discussed in [16] (in the context of Littlewood-Paley theory).

We close this section by returning full-circle to the notion of pointwise convergence. The following theorem was proved by Carleson for $L^2$-functions and by Hunt for $L^p$-functions ($p \neq 2$).

**Theorem 9** If $f$ is an $L^p$-function for $p > 1$, then its Fourier series converges to it at almost all points.

By *almost all points*, we mean that the set of points where divergence occurs has Lebesgue measure zero. References for the proof of Theorem 9 can be found in [7] and [16]. Its proof is undoubtedly the most difficult one in the theory of Fourier series.

**VI. Summability of Fourier series**

In the previous sections, we noted some problems with convergence of Fourier series partial sums. Some of these problems include Kolmogorov’s example of a Fourier series for an $L^1$-function that diverges everywhere, and Gibbs’ phenomenon and ringing in the Fourier series partial sums for discontinuous functions. Another problem is Du Bois Reymond’s example of a continuous function whose Fourier series diverges on a countably infinite set of points (see [12]). It turns out
that all of these difficulties simply disappear when new summation methods, based on appropriate modifications of the partial sums, are used.

The simplest modification of partial sums, and one of the first historically to be used, is to take arithmetic means. Define the $N^{th}$ arithmetic mean $\sigma_N$ by $\sigma_N = (S_0 + S_1 + \ldots + S_{N-1})/N$. From which it follows that

$$\sigma_N(x) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) c_n e^{inx}.$$  \hspace{1cm} (27)

The factors $\left(1 - \frac{|n|}{N}\right)$ are called convergence factors. They modify the Fourier coefficients $c_n$ so that the amplitude of the higher frequency terms (for $|n|$ near $N$) are damped down towards zero. This produces a great improvement in convergence properties as shown by the following theorem.

**Theorem 10** Let $f$ be a periodic function. If $f$ is an $L^p$-function for $p \geq 1$, then $\sigma_N \to f$ in $L^p$-norm as $N \to \infty$. If $f$ is a continuous function, then $\sigma_N \to f$ uniformly as $N \to \infty$.

Notice that $L^1$-convergence is included in Theorem 10. Even for Kolmogorov’s function, it is the case that $\|f - \sigma_N\|_1 \to 0$ as $N \to \infty$. It also should be noted that no assumption, other than continuity of the periodic function, is needed in order to ensure uniform convergence of its arithmetic means.

For a proof of Theorem 10, see [7]. The key to the proof is Fejér’s integral form for $\sigma_N$:

$$\sigma_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - t) F_N(t) \, dt$$  \hspace{1cm} (28)

where Fejér’s kernel $F_N$ is defined by

$$F_N(t) = \frac{1}{N} \left(\frac{\sin Nt/2}{\sin t/2}\right)^2.$$  \hspace{1cm} (29)

In Fig. 7 we show the graph of $F_{20}$. Compare this graph with the one of Dirichlet’s kernel $D_{20}$ in Fig. 6. Unlike Dirichlet’s kernel, Fejér’s kernel is positive [$F_N(t) \geq 0$], and is close to 0 away from the origin. These two facts are the main reasons that Theorem 10 holds. The fact that Fejér’s kernel satisfies

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F_N(t) \, dt = 1$$

is also used in the proof.

An attractive feature of arithmetic means is that Gibbs’ phenomenon and ringing do not occur. For example, in Fig. 8 we show $\sigma_{100}$ for the square wave and it is
plain that these two defects are absent. For the square wave function \( f_1 \), Eq. (28) reduces to
\[
\sigma_N(x) = \frac{1}{2\pi} \int_{x-\pi/2}^{x+\pi/2} F_N(t) \, dt.
\]
As \( x \) ranges from \(-\pi\) to \( \pi \), this formula shows that \( \sigma_N(x) \) is proportional to the area of the positive function \( F_N \) over an interval of length \( \pi \) centered at \( x \). By examining Fig. 7, which is a typical graph for \( F_N \), it is easy to see why ringing and Gibbs’ phenomenon do not occur for the arithmetic means of the square wave.

The method of arithmetic means is just one example from a wide range of summation methods for Fourier series. These summation methods are one of the major elements in the area of finite impulse response filtering in the fields of electrical engineering and signal processing.

A summation kernel \( K_N \) is defined by
\[
K_N(x) = \sum_{n=-N}^{N} m_n e^{inx}.
\] (30)
The real numbers \( \{m_n\} \) are the convergence factors for the kernel. We have already seen two examples: Dirichlet’s kernel (where \( m_n = 1 \)) and Fejér’s kernel (where \( m_n = 1 - |n|/N \)).
When $K_N$ is a summation kernel, then we define the modified partial sum of $f$ to be $\sum_{n=-N}^{N} m_n c_n e^{i n x}$. It then follows from (14) and (30) that

$$\sum_{n=-N}^{N} m_n c_n e^{i n x} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - t) K_N(t) \, dt.$$  \hfill (31)

The function defined by both sides of Eq. (31) is denoted by $K_N * f$. It is usually more convenient to use the left side of (31) to compute $K_N * f$, while for theoretical purposes (such as proving Theorem 11 below), it is more convenient to use the right side of (31).

We say that a summation kernel $K_N$ is regular if it satisfies the following three conditions.

1. For each $N$,
   $$\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(x) \, dx = 1.$$

2. There is a positive constant $C$ such that
   $$\frac{1}{2\pi} \int_{-\pi}^{\pi} |K_N(x)| \, dx \leq C.$$
3. For each $0 < \delta < \pi$,

$$\lim_{N \to \infty} \left\{ \max_{0 \leq |x| \leq \pi} |K_N(x)| \right\} = 0.$$  

There are many examples of regular summation kernels. Fejér’s kernel, which has $m_n = 1 - |n|/N$, is regular. Another regular summation kernel is Hann’s kernel, which has $m_n = 0.5 + 0.5 \cos(n\pi/N)$. A third regular summation kernel is de le Vallée Poussin’s kernel, for which $m_n = 1$ when $|n| \leq N/2$, and $m_n = 2(1-|m/N|)$ when $N/2 < |m| \leq N$. The proofs that these summation kernels are regular are given in [13]. It should be noted that Dirichlet’s kernel is not regular, because properties 2 and 3 do not hold.

As with Fejér’s kernel, all regular summation kernels significantly improve the convergence of Fourier series. In fact, the following theorem generalizes Theorem 10.

**Theorem 11** Let $f$ be a periodic function, and let $K_N$ be a regular summation kernel. If $f$ is an $L^p$-function for $p \geq 1$, then $K_N * f \to f$ in $L^p$-norm as $N \to \infty$. If $f$ is a continuous function, then $K_N * f \to f$ uniformly as $N \to \infty$.

For an elegant proof of this theorem, see [7].

From Theorem 11 we might be tempted to conclude that the convergence properties of regular summation kernels are all the same. They do differ, however, in the rates at which they converge. For example, in Fig. 9 we show $K_{100} * f_1$ where the kernel is Hann’s kernel and $f_1$ is the square wave. Notice that this graph is a much better approximation of a square wave than the arithmetic mean graph in Fig. 8. An oscilloscope, for example, can easily generate the graph in Fig. 9, thereby producing an acceptable version of a square wave.

Summation of Fourier series is discussed further in [7], [13], [15], and [16].

**VII. Generalized Fourier series**

The classical theory of Fourier series has undergone extensive generalizations during the last two hundred years. For example, Fourier series can be viewed as one aspect of a general theory of orthogonal series expansions. In this section, we shall discuss a few of the more celebrated orthogonal series, such as Legendre series, Haar series, and wavelet series.

We begin with Legendre series. The first two Legendre polynomials are defined to be $P_0(x) = 1$, and $P_1(x) = x$. For $n = 2, 3, 4, \ldots$, the $n$th Legendre polynomial $P_n$ is defined by the recursion relation

$$nP_n(x) = (2n - 1)xP_{n-1}(x) + (n - 1)P_{n-2}(x).$$
These polynomials satisfy the following orthogonality relation
\[
\int_{-1}^{1} P_n(x) P_m(x) \, dx = \begin{cases} 
0 & \text{if } m \neq n \\
\frac{(2n+1)/2}{2n+1} & \text{if } m = n.
\end{cases}
\] (32)

This equation is quite similar to Eq. (14). Because of (32)—recall how we used (14) to derive (9)—the Legendre series for a function \( f \) over the interval \([-1, 1]\) is defined to be
\[
\sum_{n=0}^{\infty} c_n P_n(x)
\]
where
\[
c_n = \frac{2}{2n+1} \int_{-1}^{1} f(x) P_n(x) \, dx.
\] (34)

The partial sum \( S_N \) of the series in (33) is defined to be
\[
S_N(x) = \sum_{n=0}^{N} c_n P_n(x).
\]

As an example, let \( f(x) = 1 \) for \( 0 \leq x \leq 1 \) and \( f(x) = 0 \) for \(-1 \leq x < 0\). The Legendre series for this step function is (see [12]):
\[
\frac{1}{2} + \sum_{k=0}^{\infty} \frac{(-1)^k(4k+3)(2k)!}{4k+1(k+1)k!} P_{2k+1}(x).
\]
In Fig. 10 we show the partial sum $S_{11}$ for this series. The graph of $S_{11}$ is reminiscent of a Fourier series partial sum for a step function. In fact, the following theorem is true.

**Theorem 12** If $\int_{-1}^{1} |f(x)|^2 \, dx$ is finite, then the partial sums $S_N$ for the Legendre series for $f$ satisfy

$$\lim_{N\to\infty} \int_{-1}^{1} |f(x) - S_N(x)|^2 \, dx = 0.$$ 

Moreover, if $f$ is Lipschitz at a point $x_0$, then $S_N(x_0) \to f(x_0)$ as $N \to \infty$.

This theorem is proved in [15] (see also [6]). Further details and other examples of orthogonal polynomial series can be found in either [3], [6], or [15]. There are many important orthogonal series—such as Hermite, Laguerre, and Tchebysheff—which we cannot examine here because of space limitations.

We now turn to another type of orthogonal series, the Haar series. The defects, such as Gibbs’ phenomenon and ringing, that occur with Fourier series expansions can be traced to the unlocalized nature of the functions used for expansions. The complex exponentials used in classical Fourier series, and the polynomials used in Legendre series, are all non-zero (except possibly for a finite number of points)
Fourier series

over their domains. In contrast, Haar series make use of localized functions, which are non-zero only over tiny regions within their domains.

In order to define Haar series, we first define the fundamental Haar wavelet $H(x)$ by

$$H(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1/2 \\ -1 & \text{if } 1/2 \leq x \leq 1. \end{cases}$$

The Haar wavelets $\{H_{j,k}(x)\}$ are then defined by

$$H_{j,k}(x) = 2^{j/2}H(2^j x - k)$$

for $j = 0, 1, 2, \ldots; k = 0, 1, \ldots, 2^j - 1$. Notice that $H_{j,k}(x)$ is non-zero only on the interval $[k2^{-j}, (k + 1)2^{-j}]$, which for large $j$ is a tiny subinterval of $[0, 1]$. As $k$ ranges between 0 and $2^j - 1$, these subintervals partition the interval $[0, 1]$, and the partition becomes finer (shorter subintervals) with increasing $j$.

The Haar series for a function $f$ is defined by

$$b + \sum_{j=0}^{\infty} \sum_{k=0}^{2^{j-1}} c_{j,k} H_{j,k}(x)$$

(35)

with $b = \int_0^1 f(x) \, dx$ and

$$c_{j,k} = \int_0^1 f(x) H_{j,k}(x) \, dx.$$

The definitions of $b$ and $c_{j,k}$ are justified by orthogonality relations between the Haar functions (similar to the orthogonality relations that we used above to justify Fourier series and Legendre series).

A partial sum $S_N$ for the Haar series in Eq. (35) is defined by

$$S_N(x) = b + \sum_{\{j,k\mid 2^j + k \leq N\}} c_{j,k} H_{j,k}(x).$$

For example, let $f$ be the function on $[0, 1]$ defined as follows

$$f(x) = \begin{cases} x - 1/2 & \text{if } 1/4 < x < 3/4 \\ 0 & \text{if } x \leq 1/4 \text{ or } 3/4 \leq x. \end{cases}$$

In Fig. 11 we show the Haar series partial sum $S_{256}$ for this function. Notice that there is no Gibbs’ phenomenon with this partial sum. This contrasts sharply with the Fourier series partial sum, also using 257 terms, which we show in Fig. 12.

The Haar series partial sums satisfy the following theorem (proved in [2] and in [10]).
Figure 11: Haar series partial sum $S_{256}$, which has 257 terms.

**Theorem 13** Suppose that $\int_0^1 |f(x)|^p \, dx$ is finite, for $p \geq 1$. Then the Haar series partial sums for $f$ satisfy

$$\lim_{N \to \infty} \left[ \int_0^1 |f(x) - S_N(x)|^p \, dx \right]^{1/p} = 0.$$ 

If $f$ is continuous on $[0, 1]$, then $S_N$ converges uniformly to $f$ on $[0, 1]$.

This theorem is reminiscent of Theorems 10 and 11 for the modified Fourier series partial sums obtained by arithmetic means or by a regular summation kernel. The difference here, however, is that for the Haar series no modifications of the partial sums are needed.

One glaring defect of Haar series is that the partial sums are discontinuous functions. This defect is remedied by the wavelet series discovered by Meyer, Daubechies, and others. The fundamental Haar wavelet is replaced by some new fundamental wavelet $\Psi$ and the set of wavelets $\{\psi_{j,k}\}$ is then defined by $\psi_{j,k}(x) = 2^{-j/2} \psi[2^j x - k]$. (Note: the bracket symbolism $\psi[2^j x - k]$ means that the value, $2^j x - k \mod 1$, is evaluated by $\psi$. This technicality is needed in order to ensure periodicity of $\psi_{j,k}$.) For example, in Fig. 13, we show graphs of $\psi_{4,1}$ and $\psi_{6,46}$ for one of the Daubechies wavelets (a Coif18 wavelet)—which is continuously differentiable. For a complete discussion of the definition of these wavelet functions, see [2] or [8].
The wavelet series, generated by the fundamental wavelet $\Psi$, is defined by

$$
N \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} \Psi_{j,k}(x)
$$

with $b = \int_{0}^{1} f(x) \, dx$ and

$$
c_{j,k} = \int_{0}^{1} f(x) \, \Psi_{j,k}(x) \, dx.
$$

This wavelet series has partial sums $S_N$ defined by

$$
S_N(x) = b + \sum_{\{j,k|2^j+k \leq N\}} c_{j,k} \Psi_{j,k}(x).
$$

Notice that when $\Psi$ is continuously differentiable, then so is each partial sum $S_N$. These wavelet series partial sums satisfy the following theorem, which generalizes Theorem 13 for Haar series (for a proof, see [2] or [10]).

**Theorem 14** Suppose that $\int_{0}^{1} |f(x)|^p \, dx$ is finite, for $p \geq 1$. Then the Daubechies wavelet series partial sums for $f$ satisfy

$$
\lim_{N \to \infty} \left[ \int_{0}^{1} |f(x) - S_N(x)|^p \, dx \right]^{1/p} = 0.
$$

If $f$ is continuous on $[0, 1]$, then $S_N$ converges uniformly to $f$ on $[0, 1]$. 

---

*Figure 12: Fourier series partial sum $S_{128}$, which has 257 terms.*
Theorem 14 does not reveal the full power of wavelet series. In almost all cases, it is possible to rearrange the terms in the wavelet series in any manner whatsoever and convergence will still hold. One reason for doing a rearrangement is in order to add the terms in the series with coefficients of largest magnitude (thus largest energy) first so as to speed up convergence to the function. Here is a convergence theorem for such permuted series:

**Theorem 15** Suppose that \( \int_{I}^{+} |f(x)|^p \, dx \) is finite, for \( p > 1 \). If the terms of a Daubechies wavelet series are permuted (in any manner whatsoever), then the partial sums \( S_N \) of the permuted series satisfy

\[
\lim_{N \to \infty} \int_{0}^{1} |f(x) - S_N(x)|^p \, dx^{1/p} = 0.
\]

If \( f \) is uniformly Lipschitz, then the partial sums \( S_N \) of the permuted series converge uniformly to \( f \).

This theorem is proved in [2] and [10]. This type of convergence of wavelet series is called **unconditional convergence**. It is known (see [8]) that unconditional convergence of wavelet series ensures an optimality of compression of signals. For details about compression of signals and other applications of wavelet series, see [14] for a simple introduction and [8] for a thorough treatment.
VIII. Discrete Fourier series

The digital computer has revolutionized the practice of science in the latter half of the twentieth century. The methods of computerized Fourier series, based upon the fast Fourier transform algorithms for digital approximation of Fourier series, have completely transformed the application of Fourier series to scientific problems. In this section, we shall briefly outline the main facts in the theory of discrete Fourier series.

The Fourier series coefficients \( c_n \) can be discretely approximated via Riemann sums for the integrals in Eq. (9). For a (large) positive integer \( M \), let \( x_k = -\pi + 2\pi k / M \) for \( k = 0, 1, 2, \ldots, M - 1 \) and let \( \Delta x = 2\pi / M \). Then the \( n \)th Fourier coefficient \( c_n \) for a function \( f \) is approximated as follows:

\[
c_n \approx \frac{1}{2\pi} \sum_{k=0}^{M-1} f(x_k) e^{-i2\pi nx_k} \Delta x = \frac{e^{-i\pi n}}{M} \sum_{k=0}^{M-1} f(x_k) e^{-i2\pi kn / M}.
\]

The last sum above is called the Discrete Fourier Transform (DFT) of the finite sequence of numbers \( \{ f(x_k) \} \). That is, we define the DFT of a sequence \( \{ g_k \} _{k=0}^{M-1} \) of numbers by

\[
G_n = \sum_{k=0}^{M-1} g_k e^{-i2\pi kn / M}.
\]

The DFT is the set of numbers \( \{ G_n \} \), and we see from the discussion above that the Fourier coefficients of a function \( f \) can be approximated by a DFT (multiplied by the factors \( e^{-i\pi n / M} \)). For example, in Fig. 14 we show a graph of approximations of the Fourier coefficients \( \{ c_n \} _{n=-50}^{50} \) of the square wave \( f_1 \) obtained via a DFT (using \( M = 1024 \)). For all values, these approximate Fourier coefficients differ from the exact coefficients by no more than \( 10^{-3} \). By taking \( M \) even larger, the error can be reduced still further.

The two principal properties of DFTs are (1) they can be inverted, (2) they preserve energy (up to a scale factor). The inversion formula for the DFT is

\[
g_k = \sum_{n=0}^{M-1} G_n e^{i2\pi kn / M}.
\]

And the conservation of energy property is

\[
\sum_{k=0}^{M-1} |g_k|^2 = \frac{1}{N} \sum_{n=0}^{M-1} |G_n|^2.
\]
Interpreting a sum of squares as energy, Eq. (40) says that, up to multiplication by the factor $1/N$, the energy of the discrete signal $\{g_k\}$ and its DFT $\{G_n\}$ are the same. These facts are proved in [1] and [13].

An application of inversion of DFTs is to the calculation of Fourier series partial sums. If we substitute $x_k = -\pi + 2\pi k/M$ into the Fourier series partial sum $S_N(x)$ we obtain (assuming that $N < M/2$ and after making a change of indices $m = n + N$):

$$S_N(x_k) = \sum_{n=-N}^{N} c_n e^{i n (-\pi+2\pi k/M)}$$

$$= \sum_{n=-N}^{N} c_n (-1)^n e^{i 2\pi n k/M}$$

$$= \sum_{m=0}^{2N} c_{m-N} (-1)^{m-N} e^{-i 2\pi kN/M} e^{i 2\pi km/M}.$$ 

Thus, if we let $g_m = c_{m-N}$ for $m = 0, 1, \ldots, 2N$ and $g_m = 0$ for $m = 2N +$
Fourier series

1, \ldots, M - 1, we have

\[ S_M(x_k) = e^{-i2\pi kN/M} \sum_{m=0}^{M-1} g_m (-1)^{m-N} e^{i2\pi km/M}. \]

This equation shows that \( S_M(x_k) \) can be computed using a DFT inversion (along with multiplications by exponential factors). By combining DFT approximations of Fourier coefficients with this last equation, it is also possible to approximate Fourier series partial sums, or arithmetic means, or other modified partial sums. See [1] or [13] for further details.

These calculations with DFTs are facilitated on a computer using various algorithms which are all referred to as fast Fourier transforms (FFTs). Using FFTs, the process of computing DFTs, and hence Fourier coefficients and Fourier series, is now practically instantaneous. This allows for rapid, so-called real-time, calculation of the frequency content of signals. One of the most widely used applications is in calculating spectrograms. A spectrogram is calculated by dividing a signal (typically a recorded, digitally sampled, audio signal) into a successive series of short duration subsignals, and performing an FFT on each subsignal. This gives a portrait of the main frequencies present in the signal as time proceeds. For example, in Fig. 15(a) we analyze discrete samples of the function

\[
\sin(2\nu_1 \pi x)e^{-100\pi(x-0.2)^2} + [\sin(2\nu_1 \pi x) + \cos(2\nu_2 \pi x)] e^{-50\pi(x-0.5)^2} + \sin(2\nu_2 \pi x)e^{-100\pi(x-0.8)^2} \tag{41}
\]

where the frequencies \( \nu_1 \) and \( \nu_2 \) of the sinusoidal factors are 128 and 256, respectively. The signal is graphed at the bottom of Fig. 15(a) and the magnitudes of the values of its spectrogram are graphed at the top. The more intense spectrogram magnitudes are shaded more darkly, while white regions indicate magnitudes that are essentially zero. The dark blobs in the graph of the spectrogram magnitudes clearly correspond to the regions of highest energy in the signal and are centered on the frequencies 128 and 256, the two frequencies used in (41).

As a second example, we show in Fig. 15(b) the spectrogram magnitudes for the signal

\[ e^{-5\pi[(x-0.5)/0.4]^{10}} [\sin(400\pi x^2) + \sin(200\pi x^2)] \tag{42} \]

This signal is a combination of two tones with sharply increasing frequency of oscillations. When run through a sound generator, it produces a sharply rising pitch. Signals like this bear some similarity to certain bird calls, and are also used in radar. The spectrogram magnitudes for this signal are shown in Fig. 15(b). We
can see two, somewhat blurred, line segments corresponding to the factors $400\pi x$ and $200\pi x$ multiplying $x$ in the two sine factors in (42).

One important area of application of spectrograms is in *speech coding*. As an example, in Fig. 16 we show spectrogram magnitudes for two audio recordings. The spectrogram magnitudes in Fig. 16(a) come from a recording of a four year old girl singing the phrase “Twinkle, twinkle, little star,” and the spectrogram magnitudes in Fig. 16(b) come from a recording of the author of this article singing the same phrase. The main frequencies are seen to be in harmonic progression (integer multiples of a lowest, fundamental frequency) in both cases, but the young girl’s main frequencies are higher (higher in pitch) than the adult male’s. The slightly curved ribbons of frequency content are known as *formants* in linguistics. For more details on the use of spectrograms in signal analysis, see [8].

It is possible to invert spectrograms. In other words, we can recover the original signal by inverting the succession of DFTs that make up its spectrogram. One application of this inverse procedure is to the *compression of audio signals*. After discarding (setting to zero) all the values in the spectrogram with magnitudes below a threshold value, the inverse procedure creates an approximation to the signal which uses significantly less data than the original signal. For example, by dis-
carding all of the spectrogram values having magnitudes less than $1/320$ times the largest magnitude spectrogram value, the young girl’s version of “Twinkle, twinkle, little star” can be approximated, without noticeable degradation of quality, using about one-eighth the amount of data as the original recording. Some of the best results in audio compression are based on sophisticated generalizations of this spectrogram technique—referred to either as lapped transforms or as local cosine expansions (see [9] and [8]).

IX. Conclusion

In this article, we have outlined the main features of the theory and application of one-variable Fourier series. Much additional information, however, can be found in the references. In particular, we did not have sufficient space to discuss the intricacies of multi-variable Fourier series—which, for example, have important applications in crystallography and molecular structure determination. For a mathematical introduction to multi-variable Fourier series, see [7], and for an introduction to their applications, see [12].

Figure 16: Spectrograms of audio signals. (a) Bottom graph displays data from a recording of a young girl singing “Twinkle, twinkle, little star.” Top graph displays the spectrogram magnitudes for this recording. (b) Similar graphs for the author’s rendition of “Twinkle, twinkle, little star.”
References


