Fourier Theory
& The Application to The Heat Equation

Julian Schennach

June 8, 2025

1 Introduction

In many applications, closed-form analytic solutions are not always available. Fortunately, there exist many approximization methods, including the Weierstraß Approximation Theorem, for instance, which guarantee that any continuous function can be approximated arbitrarily closely by a sum of polynomials (see my expository paper from Measure Theory). A similar result, implied by the Stone-Weierstraß Theorem, states that continuous (and sometimes even discontinuous) functions can be approximated arbitrarily closely by so-called trigonometric polynomials. In the limit of an infinite number of terms, we obtain the Fourier series, named after the physicist Joseph Fourier:

Definition 1. The *Fourier series* of a function f of period P that is integrable over $[0, P]$ is defined as $\sum_{k=-\infty}^{\infty} a_k e^{i2\pi \frac{k}{P} x}$, where the coefficients a_k can be proven to be $\frac{1}{P} \int_0^P f(t) e^{-i2\pi \frac{k}{P} t}\, dt$.

Unfortunately, this series does not always converge to f, as will be demonstrated with a counterexample. For the sake of simplicity, we will only be referring to a condition for convergence based on the notion of "bounded variation" that guarantees (but in fact is not necessary for) pointwise convergence. While we can only create a Fourier series for a *periodic* function, there does exist a generalization of the Fourier coefficients to a set of non-periodic functions:

Definition 2. Assume that a given function f Lebesgue-integrable and absolutely integrable, meaning that $\int_{-\infty}^{\infty} |f(x)|\, dx$ is finite. Then, the *Fourier transform* of f is defined as:

$$(\mathcal{F}f)(x) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi x t}\, dt.$$

This new function represents a *frequency distribution.*

The aforementioned Fourier series decomposes a function f into a sum of complex exponentials of different periods, and so it should not be surprising that we are able to use the Fourier transform to represent f, not as a sum, but as an integral of these complex exponentials:

Theorem 1. (Fourier Inversion Theorem)

If a Lebesgue-integral continuous function f as the Fourier transform $\mathcal{F}f$, then:

$$f(x) = \int_{-\infty}^{\infty} (\mathcal{F}f)(t) e^{i2\pi xt}\, dt.$$

In other words, if we have a Fourier transform $\mathcal{F}f$, we can reconstruct the original function f — in fact, $f(x)$ is the Fourier transform of $\mathcal{F}f(-x)$. Of course, this continuous Fourier transform may not always be very practical and our original Fourier series may be more desirable — indeed, if we have some computer program that calculates the Fourier transform of a function (e.g. for the processing of images), we cannot store the transform for each of the infinitely many possible periods.

In addition to presenting these definitions for the Fourier series and transform and proving the Fourier Inversion Theorem, in this paper we will also be proving Perseval's Theorem — which states that the Fourier transform is a unitary operator on $L^2(\mathbb{R})$ — and the Poisson Summation Formula — which relates the periodic sum of a function to the same periodic sum of its Fourier transform. At the end, we will be analyzing a simple but famous application of Fourier analysis in physics, Fourier's heat equation.

2 The Fourier Series & Convergence Problems

Select a complex-valued function f with period P such that f can be expressed as $\sum_{k=-\infty}^{\infty} a_k e^{i2\pi \frac{k}{P} x}$, where the division by P in the exponential guarantees that $f(x) = f(x+P)$. By Euler's identity, we can rewrite the Fourier series as a trigonometric polynomial of infinite degree, $\sum_{k=-\infty}^{\infty}(A_k \cos(2\pi \frac{k}{P} x) + B_k \sin(2\pi \frac{k}{P} x))$, for appropriate coefficients $A_k = a_k, B_k = ia_k$.

In order to derive the coefficients a_k, we assume that $\sum_{-\infty}^{\infty} |a_k|$ is finite (the expressions for the coefficients also work if this sum diverges). We demonstrate that the coefficients a_k are of the form $\frac{1}{P}\int_0^P f(t)e^{-i2\pi \frac{k}{P} x}\, dt$, as suggested above. We first multiply both sides of the assumed equality $f(x) = \sum_{k=-\infty}^{\infty} a_k e^{i2\pi \frac{k}{P} x}$ by $e^{-i2\pi \frac{n}{P} x}$ to obtain $f(x)e^{-i2\pi \frac{n}{P} x} = \sum_{k=-\infty}^{\infty} a_k e^{i2\pi \frac{k-n}{P} x}$, and then we integrate both sides from 0 to P to have $\int_0^P f(x)e^{-i2\pi \frac{n}{P} x}\, dx = \int_0^P (\sum_{k=-\infty}^{\infty} a_k e^{i2\pi \frac{k-n}{P} x})\, dx$. We express the series as the limit of a summation, making the right-hand side $\int_0^P (\lim_{\ell\to\infty} \sum_{k=-\ell}^{\ell} a_k e^{i2\pi \frac{k-n}{P} x})\, dx$. Since the magnitude of the summation is bounded by $|\sum_{k=-\ell}^{\ell} a_k e^{i2\pi \frac{k-n}{P} x}| \leq \sum_{k=-\ell}^{\ell} |a_k||e^{i2\pi \frac{k-n}{P} x}| = \sum_{k=-\ell}^{\ell} |a_k| \leq \sum_{-\infty}^{\infty} |a_k|$, a finite series. By the dominated convergence theorem, we can interchange the limit and the integral to obtain the equation:

$$\int_0^P f(x)e^{-i2\pi \frac{n}{P} x}\, dx = \lim_{\ell\to\infty} \int_0^P (\sum_{k=-\ell}^{\ell} a_k e^{i2\pi \frac{k-n}{P} x})\, dx = \sum_{k=-\infty}^{\infty} \int_0^P a_k e^{i2\pi \frac{k-n}{P} x}\, dx.$$

Since $\int_0^P a_n e^{i2\pi \frac{n-n}{P} x}\, dx = \int_0^P a_n\, dx = Pa_n$ is the only nonzero integral, with all others being $\int_0^P a_k e^{i2\pi \frac{k-n}{P} x}\, dx = [\frac{P}{2i\pi} a_k e^{i2\pi \frac{k-n}{P} x}]_0^P = 0$, we determine that $\int_0^P f(x)e^{-i2\pi \frac{n}{P} x}\, dx = Pa_n$. We have thus proven that, if f is writable as a sum of complex exponentials, then the coefficients are $a_k = \frac{1}{P}\int_0^P f(t)e^{-i2\pi \frac{k}{P} x}\, dt$.

The fact that we integrate f over the whole period in order to determine the coefficients already indicates that the Fourier series does not always accurately represent f. Instead, the coefficients only represent the "average" contribution of each exponential, leading to the possibility that the Fourier series does not converge to f. More explicitly, if $f_n(x) = \sum_{k=-n}^{n} a_k e^{i2\pi \frac{k}{P} x}$ is the truncation of the Fourier series, it is not always the case that $\lim_{n\to\infty} f_n(x) = f(x)$ (i.e. there does not necessarily exist pointwise convergence).

Example 1. (Square wave approximation)

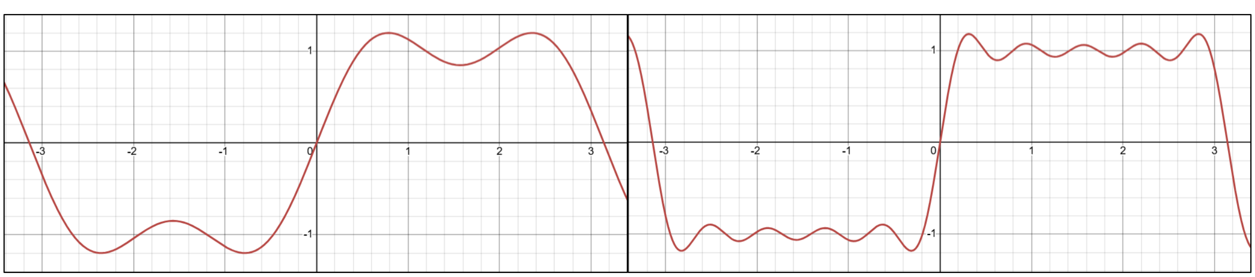
Define the discontinuous but periodic function $g(x) = \{ \begin{smallmatrix} 1 & x\in[0,\pi) \\ -1 & x\in[\pi,2\pi) \end{smallmatrix}$ such that $g(x) = g(x + 2\pi)$. We determine the Fourier coefficients:

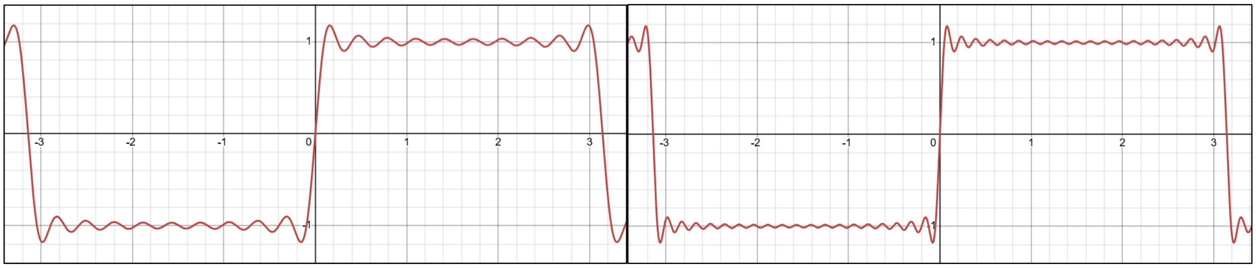
$$\begin{aligned} a_k = \frac{1}{2\pi}\int_0^{2\pi} g(t)e^{-ikx}\,dt &= \frac{1}{2\pi}\left(\int_0^{\pi} e^{-ikx}\,dt - \int_{\pi}^{2\pi} e^{-ikx}\,dt\right) \\ &= \frac{1}{2\pi}(-[\frac{1}{ik}e^{-ikx}]_0^{\pi} + [\frac{1}{ik}e^{-ikx}]_{\pi}^{2\pi}) \\ &= \frac{1}{ik\pi} - \frac{1}{ik\pi}e^{-ik\pi} \end{aligned}$$

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} g(t)\,dt = 0$$

$$\begin{aligned} a_{-k} = \frac{1}{2\pi}\int_0^{2\pi} g(t)e^{ikx}\,dt &= \frac{1}{2\pi}\left(\int_0^{\pi} e^{ikx}\,dt - \int_{\pi}^{2\pi} e^{ikx}\,dt\right) \\ &= \frac{1}{2\pi}([\frac{1}{ik}e^{ikx}]_0^{\pi} - [\frac{1}{ik}e^{ikx}]_{\pi}^{2\pi}) \\ &= -\frac{1}{ik\pi} + \frac{1}{ik\pi}e^{ik\pi}. \end{aligned}$$

The coefficients a_k and a_{-k} are only nonzero for odd k, so we obtain the truncations $g_n(x) = \sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \frac{2}{i(2\ell-1)\pi}(e^{i(2\ell-1)x} - e^{-i(2\ell-1)x}) = \frac{4}{\pi}\sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{2\ell-1}\sin((2\ell-1)x)$, which are plotted below for $n = 2, 5, 10, 20$. Clearly, at $x = 2\pi m$ for integer m, all truncations evaluate to $g_n(2\pi m) = 0$, implying that the limit is 0, even if the point on the function being approximated is $g(2\pi n) = 1$. Since the continuous functions that are the Fourier series truncations must approximate a discontinuous function, it is to be expected that the approximations approach the average value of the function at the points of discontinuity rather than its true value.





Example 2. (Sawtooth wave approximation)

The sawtooth wave, defined as $h(x) = x$ for $x \in [-\pi, \pi)$ such that $h(x) = h(x + 2\pi)$, has the Fourier series truncations $h_n(x) = -\frac{2}{\pi} \sum_{k=1}^{n} \frac{(-1)^k}{k} \sin(kx)$, again plotted for $n = 2, 5, 10, 20$. Clearly, $h_n(\pi) = 0$ again, resulting in the Fourier series being 0, but $h(\pi) = -\pi$.

Fortunately, as can be seen in the graphs above, at all other x, we indeed have $\lim_{n\to\infty} g_n(x) = g(x)$ and $\lim_{n\to\infty} h_n(x) = h(x)$. We will soon prove these results when considering the pointwise convergence conditions. Another glaring issue with the successive approximations of the square wave in particular (and, less noticeably, of the sawtooth wave) is the fact that, close to the points of discontinuity, there exists one oscillation that deviates extremely from the actual function. Unfortunately, the Gibbs phenomenon [2] is not a byproduct of technical limitations, as the physicist A. A. Michelson allegedly suspected when he constructed a machine that outputs Fourier series approximations. Worse yet, the fluctuation's amplitude never decreases to 0 — only its width does!

Proof. (Limit of the amplitude)

At the local maxima, we require that $g_n'(x) = \frac{4}{\pi} \sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{2\ell-1} \sin'((2\ell-1)x) = \frac{4}{\pi} \sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \cos((2\ell-1)x) = 0$. The crest of the deviating oscillation occurs at the smallest positive solution, specifically $x = \frac{\pi}{2\lfloor \frac{n}{2} \rfloor}$:

$$\frac{4}{\pi} \sum_{\ell=1}^{\lfloor \frac{n}{2} \rfloor} \cos\left(\frac{(2\ell-1)\pi}{2\lfloor \frac{n}{2} \rfloor}\right) = \frac{4}{\pi} \sum_{\ell=1}^{\lfloor \frac{\lfloor \frac{n}{2} \rfloor}{2} \rfloor} \left(\cos\left(\frac{(2\ell-1)\pi}{2\lfloor \frac{n}{2} \rfloor}\right) + \cos\left(\pi - \frac{(2\ell-1)\pi}{2\lfloor \frac{n}{2} \rfloor}\right)\right) = 0.$$

For any smaller positive x, there exist too few negative terms of too small magnitude for the sum to be cancel to 0. Evaluated at this maximizing x, the summation in $g_n(x)$ approximates an integral in the limit:

$$\lim_{n\to\infty} g_n\left(\frac{\pi}{2\lfloor\frac{n}{2}\rfloor}\right) = \frac{2}{\pi}\lim_{n\to\infty}\sum_{\ell=1}^{\lfloor\frac{n}{2}\rfloor}\frac{2\lfloor\frac{n}{2}\rfloor}{2\ell-1}\sin\left(\frac{(2\ell-1)\pi}{2\lfloor\frac{n}{2}\rfloor}\right)\frac{1}{\lfloor\frac{n}{2}\rfloor} = \frac{2}{\pi}\int_0^{\pi}\frac{1}{t}\sin t\,dt.$$

Unfortunately, the sinc function ($\frac{\sin x}{x}$) has no analytical antiderivative, so we need to calculate the integral numerically. We then find that the deviation from the square wave is approximately $-1+\frac{2}{\pi}\int_0^{\pi}\frac{1}{t}\sin t\,dt \approx 0.1789797\cdots$, which is rather significant. The value of the definite integral is known as the Wilbraham-Gibbs constant, named after the discoverer of the phenomenon, the mathematician Henry Wilbraham (the reference to the physicist Josiah Gibbs is another example of Stigler's law).

□

This $\sim 17.898\%$ error generalizes to any discontinuous function. The Gibbs phenomenon does not pose trouble for pointwise convergence, but does for uniform convergence: For $\epsilon > 0$ less than ~ 0.17898, there exists no sufficiently large N such that $|g_n(x) - g(x)| < \epsilon$ for all $n > N$ and for all x. For the sake of simplicity, we will only show the sufficient conditions for the Fourier series approximations of a function to converge *pointwise*, but this brief discussion demonstrates once more the distinction between uniform and pointwise convergence.

Definition 3. A function is of *bounded variation* over an interval $[a, b]$ if

$$\sup_{n, a\le x_0\le x_1\le\cdots\le x_n\le b}\sum_{i=0}^{n-1}|f(x_{i+1}) - f(x_i)|$$

is finite. Equivalently, f must be expressible as the difference of two monotonically increasing functions.

Theorem 2. (Dirichlet-Jordan criterion)

If the integrable function f of period P is of bounded variation over $[-P, P]$, then the limit of the Fourier series truncations, $\lim_{n\to\infty} f_n(x)$, is equal to $\lim_{\epsilon\to 0}\frac{f(x+\epsilon)+f(x-\epsilon)}{2}$, which takes into consideration discontinuities of f. When f is continuous at x, we have $\lim_{n\to\infty} f_n(x) = f(x)$.

The square and sawtooth waves satisfy the Dirichlet-Jordan criterion because the suprema are 2 and 4π, respectively, and so do all smooth functions when restricted to the domain $[-P, P]$. Do note that functions that do not satisfy the criterion may still be equal to their Fourier series.

The proof for this convergence test exploits the properties of the Dirichlet kernel, defined as $D_n(t) = \frac{\sin(\frac{(2n+1)\pi t}{P})}{2\sin(\frac{\pi t}{P})}$, where P is again the period of f. It turns out that the truncated Fourier series $\sum_{k=-n}^{n} a_k e^{i2\pi\frac{k}{P}x}$ is equal to $\frac{1}{P}\int_{-P}^{P} f(x-t)D_n(t)\,dt$. Unfortunately, we require a significant number of lemmas, and in order to not pursue a tangent, we leave the proof as an exercise for the reader.

3 The Fourier Transform & Its Properties

We now wish to generalize the above results of Fourier series to non-periodic functions. Recall the definition from the introduction of the Fourier transform $\mathcal{F}f$ for a Lebesgue-integrable function f, which is also be expressed as $\hat{f}$ or (rarely) $\overline{f}$ in different contexts. The notation we choose clearly indicates that the Fourier transform is an operator, a function on a set of function, just like the familiar differential operator D. Of course, $\mathcal{F}$ is very distinct from D, most notably that $\mathcal{F}$ can map real functions to complex ones, while D never does. For instance, if g is defined as the (real) indicator function for the set $[0, 1]$ (i.e. $g(x) = 1$ when x is in the set, 0 otherwise), we have $(\mathcal{F}g)(x) = \int_{-\infty}^{\infty} g(t)e^{-i2\pi xt}\,dt = \int_0^1 e^{-i2\pi xt}\,dt = [-\frac{1}{i2\pi}e^{-i2\pi xt}]_0^1 = \frac{1-e^{i2\pi x}}{i2\pi}$, which is a complex function.

Nevertheless, some of the properties of D do transfer over to $\mathcal{F}$, perhaps most obviously linearity. Fortunately, by the linearity of the integral, we can verify that the Fourier transform $(\mathcal{F}(af + bg))(x)$ is equal to:

$$\int_{-\infty}^{\infty}(af(t)+bg(t))e^{-i2\pi xt}\,dt = a\int_{-\infty}^{\infty} f(t)e^{-i2\pi xt}\,dt+b\int_{-\infty}^{\infty} g(t)e^{-i2\pi xt}\,dt = a(\mathcal{F}f)(x)+b(\mathcal{F}g)(x).$$

Let $a(x) = ax$. We know that $D(f \circ a)(x) = a(Df)(ax)$ through the chain rule, and $\mathcal{F}$ functions very similarly, as can be proven through a straightforward change of variables:

$$(\mathcal{F}(f \circ a))(x) = \int_{-\infty}^{\infty} f(at)e^{-i2\pi xt}\,dt = \frac{1}{a}\int_{-\infty}^{\infty} f(s)e^{-i2\pi\frac{xs}{a}}\,ds = \frac{1}{a}(\mathcal{F}f)\left(\frac{x}{a}\right).$$

The operator D is famously not invertible: given the derivative Df, we cannot determine f (if the indefinite integral of Df is definable, we only obtain f up to a constant). In contrast, the Fourier transform $\mathcal{F}$ has a very elegant theorem that guarantees its invertibility. [4]

Theorem 3. (Fourier Inversion Theorem)

If a Lebesgue-integrable continuous function f has the Fourier transform $\mathcal{F}f$, then:

$$f(x) = \int_{-\infty}^{\infty}(\mathcal{F}f)(t)e^{i2\pi xt}\,dt.$$

Proof. We know that $\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(\mathcal{F}f)(t)e^{i2\pi xt}e^{-\pi\epsilon^2t^2}\,dt = \int_{-\infty}^{\infty}(\mathcal{F}f)(t)e^{i2\pi xt}(\lim_{\epsilon\to 0^+}e^{-\pi\epsilon^2t^2})\,dt = \int_{-\infty}^{\infty}(\mathcal{F}f)(t)e^{i2\pi xt}\,dt$ by the dominated convergence theorem because $\mathcal{F}f$ and $e^{-\pi\epsilon^2|t|^2}$ are both Lebesgue integrable. We now define the function $g(y) = e^{i2\pi xy-\pi\epsilon^2y^2}$, which simplifies the desired integral to $\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(\mathcal{F}f)(t)g(t)\,dt$ or $\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(\int_{-\infty}^{\infty} f(\tau)e^{-i2\pi t\tau}\,d\tau)g(t)\,dt$. Under the conditions of Fubini's Theorem (i.e. when f, g, and $e^{-i2\pi t\tau}$ are Lebesgue integrable, which is the case), we have $\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(\int_{-\infty}^{\infty} f(\tau)e^{-i2\pi t\tau}\,d\tau)g(t)\,dt = \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty} f(\tau)(\int_{-\infty}^{\infty} g(t)e^{-i2\pi t\tau}\,dt)\,d\tau$. Calculating the inner integral, which is the Fourier transform of g, we have

$$\int_{-\infty}^{\infty} g(t)e^{-i2\pi t\tau}\,dt = \int_{-\infty}^{\infty} e^{-i2\pi(\tau-x)t-\pi\epsilon^2t^2}\,dt = \int_{-\infty}^{\infty}\cos(2\pi(\tau-x)t)e^{-\pi\epsilon^2t^2}\,dt$$
$$= \frac{1}{\sqrt{\pi}\epsilon}\int_{-\infty}^{\infty}\cos(\frac{2\sqrt{\pi}(\tau-x)}{\epsilon}t')e^{-t'^2}\,dt' = \frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}.$$

The imaginary component of the complex exponential vanishes because the sine function is odd. Also, the final integral of an exponential times a cosine is a well-known result. Therefore, $\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty} f(\tau)(\int_{-\infty}^{\infty} g(t)e^{-i2\pi t\tau}\,dt)\,d\tau = \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty} f(\tau)(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau$. We now need to evaluate this limit, which should ideally be equal to $f(x)$.

Being a Gaussian integral, we know that $\int_{-\infty}^{\infty}\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}\,d\tau = 1$, so:

$$(\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty} f(\tau)(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau) - f(x) = \lim_{\epsilon\to 0^+}\left(\int_{-\infty}^{\infty} f(\tau)(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau - f(x)\int_{-\infty}^{\infty}\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}\,d\tau\right) =$$

$$\lim_{\epsilon\to 0^+}\left(\int_{-\infty}^{\infty} f(\tau)(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau - \int_{-\infty}^{\infty} f(x)(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau\right) = \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(f(\tau)-f(x))(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau.$$

Since f is continuous, $\lim_{\tau\to x} f(\tau) = f(x)$, so we know that, for any sufficiently small δ, there exists some η such that $|f(\tau)-f(x)| < \eta$ when $|\tau - x| < \delta$, such that η goes to 0 when δ does. Let $\delta = \epsilon^{\frac{1}{2}}$ (so δ approaches 0 when ϵ does). We now prove that the final limit shown above tends to 0, so we can take the absolute value of the integrand to calculate an upper bound and break up the intervals integrated over:

$$\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(f(\tau)-f(x))(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau \le \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}|f(\tau)-f(x)|(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau =$$

$$\lim_{\epsilon\to 0^+}\int_{|\tau-x|<\delta}|f(\tau)-f(x)|(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau + \lim_{\epsilon\to 0^+}\int_{|\tau-x|\ge\delta}|f(\tau)-f(x)|(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau.$$

By definition, $|f(\tau)-f(x)| < \eta$ whenever $|\tau - x| < \delta$, so the first integral is clearly bounded by $\lim_{\epsilon\to 0^+}\int_{|\tau-x|<\delta}\frac{\eta}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}\,d\tau \le \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}\frac{\eta}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}\,d\tau = \lim_{\epsilon\to 0^+}\eta = 0$. Furthermore, since $\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}$ is bounded above by $\frac{1}{\epsilon}e^{-\frac{\pi\delta^2}{\epsilon^2}}$ over the interval, the second integral is bounded above by $\lim_{\epsilon\to 0^+}\int_{|\tau-x|\ge\delta}|f(\tau)-f(x)|(\frac{1}{\epsilon}e^{-\frac{\pi\delta^2}{\epsilon^2}})\,d\tau$. We defined δ to be $\epsilon^{\frac{1}{2}}$, so this bound is equal to $\lim_{\epsilon\to 0^+}\int_{|\tau-x|\ge\epsilon^{\frac{1}{2}}}|f(\tau)-f(x)|(\frac{1}{\epsilon}e^{-\pi\epsilon^{-1}})\,d\tau$. The exponential no longer depends on τ, so we can take the factor out of the integral to obtain $(\lim_{\epsilon\to 0^+}\frac{1}{\epsilon}e^{-\pi\epsilon^{-1}})(\lim_{\epsilon\to 0^+}\int_{|\tau-x|\ge\epsilon^{\frac{1}{2}}}|f(\tau)-f(x)|\,d\tau) = (\lim_{\epsilon\to 0^+}\frac{1}{\epsilon}e^{-\pi e^{-1}})(\int_{-\infty}^{\infty}|f(\tau)-f(x)|\,d\tau)$. The latter integral is finite because f is Lebesgue integrable, and $\lim_{\epsilon\to 0^+}\frac{1}{\epsilon}e^{-\pi\epsilon^{-1}}$ approaches 0 because $\lim_{\epsilon'\to\infty}\frac{\epsilon'}{e^{\pi\epsilon'}} = \lim_{\epsilon'\to\infty}\frac{1}{\pi e^{\pi\epsilon'}} = 0$ (where $\epsilon' = \frac{1}{\epsilon}$) through a simple application of l'Hopital's rule. Both integrals approach 0, so our original integral $\lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}(f(\tau)-f(x))(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau$ is 0. Thus, $\int_{-\infty}^{\infty}(\mathcal{F}f)(t)e^{i2\pi xt}\,dt = \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty} f(\tau)(\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}})\,d\tau = f(x)$, which completes the proof of the Fourier Inversion Theorem.

□

Remark 1. The function $\frac{1}{\epsilon}e^{-\frac{\pi(\tau-x)^2}{\epsilon^2}}$ converges pointwise to 0 for all $\tau \ne x$ and to ∞ for $\tau = x$ as ϵ approaches 0, and its integral over the real numbers is 1 for any ϵ. The limit is often referred to as the delta function, although it cannot be defined rigorously without the notion of distributions. The final step of the proof above is a demonstration of the sifting property of the delta function.

Notice that this inverse Fourier transform $\mathcal{F}^{-1}$ is defined very similarly to $\mathcal{F}$ — in fact, only the sign in the exponential is different. We can therefore express the Fourier Inversion Theorem as $f(-x) = (\mathcal{F}^2 f)(x)$. We thereby conclude that $\mathcal{F}^4$ is the identity operator, with $\mathcal{F}^3$ being the inverse of $\mathcal{F}$. The Fourier transform has similarities to a rotation function, not for vectors, but rather for functions. More explicitly, $\mathcal{F}$ is an example of a *unitary operator.*

Definition 4. A *Hilbert space* is a vector space imbued with a real- or complex-valued inner product $\langle \cdot, \cdot \rangle$. So, for any elements x, y, z of a Hilbert space, we require that:

- Conjugate symmetry: $\langle y, x \rangle = \overline{\langle x, y \rangle}$
- Linearity in the first argument: $\langle ax + bz, y \rangle = a\langle x, y \rangle + b\langle z, y \rangle$
- Positive definiteness $\langle x, x \rangle \geq 0$ (where the previous property implies that $\langle x, x \rangle = 0$ if and only if $x = 0$)

The Hilbert space we will be considering is $L^2(\mathbb{R})$, the set of square-integrable functions (i.e. functions f such that $\int_{-\infty}^{\infty} |f(x)|^2 \, dx$ is finite). The inner product $\langle f, g \rangle$ of two functions $f, g \in L^2(\mathbb{R})$ is defined as $\int_{-\infty}^{\infty} f(t)\overline{g(t)} \, dt$ — this integral is also known as the convolution $f * g$.

Definition 5. A *unitary operator* U is a surjective function from a Hilbert space to itself that preserves the inner product, so $\langle Uf, Ug \rangle = \langle f, g \rangle$ for any elements f, g of the Hilbert space.

Theorem 4. (Parseval's Theorem)
The Fourier transform $\mathcal{F}$ is a unitary operator on $L^2(\mathbb{R})$.

Proof. We prove that $\int_{-\infty}^{\infty} f(t)\overline{g(t)} \, dt = \int_{-\infty}^{\infty} (\mathcal{F}f)(\tau)\overline{(\mathcal{F}g)(\tau)} \, d\tau$ for any two Lebesgue-integrable functions f, g. By the Fourier Inversion Theorem, we have:

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)} \, dt = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} (\mathcal{F}f)(\tau) e^{i2\pi t\tau} \, d\tau \right) \left(\int_{-\infty}^{\infty} \overline{(\mathcal{F}g)(\tau')} e^{-i2\pi t\tau'} \, d\tau' \right) \, dt.$$

Note that the conjugation from $\overline{g(t)}$ can be distributed throughout the integral of g's Fourier transform representation. Under the assumptions of Fubini's Theorem, we can rewrite the product of integrals as the following double integral:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mathcal{F}f)(\tau)\overline{(\mathcal{F}g)(\tau')} e^{i2\pi t(\tau - \tau')} \, d\tau \, d\tau' \, dt.$$

We rearrange the order of the integrations to get:

$$\int_{-\infty}^{\infty} (\mathcal{F}f)(\tau) \left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \overline{(\mathcal{F}g)(\tau')} e^{-i2\pi t\tau'} \, d\tau' \right) e^{i2\pi t\tau} \, dt \right) \, d\tau.$$

By the Fourier Inversion Theorem, the inner double integral must be equal to simply $\overline{(\mathcal{F}g)(\tau)}$. We thus conclude that $\int_{-\infty}^{\infty} f(t)\overline{g(t)} \, dt = \int_{-\infty}^{\infty} (\mathcal{F}f)(\tau)\overline{(\mathcal{F}g)(\tau)} \, d\tau$, as desired. The fourier transform does preserve the inner product of $L^2(\mathbb{R})$

Note that the Fourier transform of any $f \in L^2(\mathbb{R})$ must be square-integrable too. In fact, $\int_\infty^\infty |f(t)|^2\, dt = \int_\infty^\infty f(t)\overline{f(t)}\, dt = \langle f, f\rangle = \langle \mathcal{F}f, \mathcal{F}f\rangle = \int_\infty^\infty |(\mathcal{F}f)(t)|^2\, dt$ (this identity is also known as the Plancherel theorem). Therefore, the Fourier transform is indeed a function from $L^2(\mathbb{R})$ to itself.

□

4 Poisson Summation Formula

So far, we have mostly analyzed the properties of the Fourier transform on one function alone. We now consider its effect on a periodic sum of the following type of function.

Definition 6. A *Schwartz function* is a smooth (infinitely differentiable) function such that $|f^{(n)}(x)| \leq c_{nN}|x|^{-N}$ for any nonnegative integer n and N.

Theorem 5. (Poisson Summation Formula)

If f is a Schwartz function, $\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{k=-\infty}^{\infty}(\mathcal{F}f)(k)e^{2\pi inx}$

Proof. The fact that f is a Schwartz function guarantees that the series $\sum_{n=-\infty}^{\infty} f(x+n)$ is a Lebesgue-integrable function, specifically with period 1. $|\sum_{n=-\infty}^{\infty} f(x+n)|$ is bounded by $\sum_{n=-\infty}^{\infty}|f(x+n)|$, which is an integrable function over the interval $[0,1]$ because it is finite over $[0,1]$. Indeed, $|f(x+n)| \leq \frac{1}{|x+n|^{n-2}} \leq \frac{1}{2^{n-2}}$ when $n \geq 2$ and $|f(x+n)| \leq \frac{1}{|x+n|^{-n-3}} \leq \frac{1}{2^{-n-3}}$ when $n \leq -3$. So, $\sum_{n=-\infty}^{\infty}|f(x+n)| \leq \sum_{n=-2}^{1}|f(x+n)| + 2\sum_{n=0}^{\infty}\frac{1}{2^n} = \sum_{n=-2}^{1}|f(x+n)| + 4$, which is a finite upper bound.

Hence, we are allowed to derive the Fourier series coefficients of the function $\sum_{n=-\infty}^{\infty} f(x+n)$. We know that $a_k = \int_0^1(\sum_{n=-\infty}^{\infty} f(t+n))e^{-i2\pi kt}\, dt = \int_0^1 \lim_{m\to\infty}(\sum_{n=-m}^{m} f(t+n)e^{-i2\pi kt})\, dt$. Since $|\sum_{n=-m}^{m} f(x+n)e^{-i2\pi kx}| \leq \sum_{n=-m}^{m}|f(x+n)e^{-i2\pi kx}| = \sum_{n=-m}^{m}|f(x+n)| \leq \sum_{n=-\infty}^{\infty}|f(x+n)|$ is an integrable function over $[0,1]$ as demonstrated above, we can interchange the integral and the limit by the dominated convergence theorem, resulting in the following expression for the coefficients:

$$a_k = \lim_{m\to\infty}\int_0^1 \sum_{n=-m}^{m} f(t+n)e^{-i2\pi kt}\, dt = \lim_{m\to\infty}\sum_{n=-m}^{m}\int_0^1 f(t+n)e^{-i2\pi kt}\, dt =$$

$$\sum_{n=-\infty}^{\infty}\int_n^{n+1} f(t)e^{i2\pi kn - i2\pi kt}\, dt = \sum_{n=-\infty}^{\infty}\int_n^{n+1} f(t)e^{-i2\pi kt}\, dt = \int_{-\infty}^{\infty} f(t)e^{-i2\pi kt}\, dt = (\mathcal{F}f)(k).$$

Note that $e^{i2\pi kn} = 1$ because k, n are integers. Therefore, the Fourier series expansion for $\sum_{n=-\infty}^{\infty} f(x+n)$ is $\sum_{k=-\infty}^{\infty}(\mathcal{F}f)(k)e^{2\pi ikx}$, and now we must only verify convergence. Since, as frequently mentioned, $\sum_{n=-\infty}^{\infty}|f(x+n)|$ is finite, we know that $(\sum_{n=-\infty}^{\infty} f(x+n))' = \sum_{n=-\infty}^{\infty} f'(x+n)$ by Tannery's theorem, implying that $\sum_{n=-\infty}^{\infty} f(x+n)$ is smooth. Thus, by the Dirichlet-Jordan criterion, the function $\sum_{n=-\infty}^{\infty} f(x+n)$ must be equal to its Fourier series $\sum_{n=-\infty}^{\infty}(\mathcal{F}f)(n)e^{2\pi inx}$.

□

Corollary 1. The Poisson summation formula is often evaluated at $x = 0$ to obtain $\sum_{n=-\infty}^{\infty} f(n) = \sum_{k=-\infty}^{\infty}(\mathcal{F}f)(k)$ — i.e. a sum of a Schwartz function over the integers is invariant under a Fourier transform (and its inverse, for that matter).

Example 3. An elegant application [3] of the Poisson summation formula is in the evaluation of $\sum_{n=-\infty}^{\infty} \frac{2c}{c^2+4\pi^2n^2}$ for complex c with positive real part. We notice that the Fourier transform of $e^{-c|x|}$ is $\int_{-\infty}^{\infty} e^{-c|t|}e^{-i2\pi xt}\,dt = \int_{-\infty}^{0} e^{ct-i2\pi xt}\,dt + \int_{0}^{\infty} e^{-ct-i2\pi xt}\,dt = [\frac{1}{c-i2\pi x}e^{ct-i2\pi xt}]_{-\infty}^{0} + [\frac{1}{-c-i2\pi x}e^{-ct-i2\pi xt}]_{0}^{\infty} = (\frac{1}{c-i2\pi x} - 0) + (0 - \frac{1}{-c-i2\pi x}) = \frac{2c}{c^2+4\pi^2n^2}$ because $\Re c > 0$. By the Poisson summation formula, $\sum_{n=-\infty}^{\infty} \frac{2c}{c^2+4\pi^2n^2} = \sum_{k=-\infty}^{\infty} e^{-c|k|}$, and since the latter series is geometric, we can calculate it explicitly as $-1+2\sum_{k=0}^{\infty} e^{-c|k|} = -1 + \frac{2}{1-e^{-c}} = \frac{1+e^{-c}}{1-e^{-c}} = \frac{e^c+1}{e^c-1}$ (i.e. $\tanh c$). We have now demonstrated that $\sum_{n=-\infty}^{\infty} \frac{2c}{c^2+4\pi^2n^2} = \frac{e^c+1}{e^c-1}$. When evaluated as $c = \pi$, we also quickly prove that $\sum_{n=1}^{\infty} \frac{1}{1+n^2} = \frac{1}{2}(-1 + \sum_{n=-\infty}^{\infty} \frac{1}{1+n^2}) = \frac{1}{2}(-1 + \pi \sum_{n=-\infty}^{\infty} \frac{4\pi}{4\pi^2+4\pi^2n^2}) = \frac{1}{2}(-1+\pi\frac{e^{2\pi}+1}{e^{2\pi}-1})$, which is far easier than the alternative derivation with a path integral around poles in the complex plane.

5 The Heat Equation — A Historical Application

We now consider the first ever application of Fourier theory [1]: In his 1822 *The Analytical Theory of Heat*, the physicist Joseph Fourier derived a differential equation that describes the evolution of the temperature $u(x, y, z, t)$ at the position x, y, z of a body over the time t.

Fourier's reasoning is based on classical and not statistical thermodynamics, so all variables are continuous and not discrete. (As a historical aside, he also relied on the obsolete caloric theory, which treated heat as a physical substance.) To derive the heat equation, let $u(x, y, z, t)$ be the temperature of a body at the position (x, y, z) at time t. At a sufficiently small local scale (in a $dx\,dy\,dz$ volume), the body is in thermal equilibrium, so we can assume that the heat transferred from one region to another, infinitessimally close one is proportional to the temperature gradient between the regions (i.e. $\vec{h}_{\text{eat transfer}}(x, y, z, t) = \kappa_1(\nabla u(x, y, z, t))$ for a positive constant κ_1). Furthermore, the change in temperature over time in this local region is proportional to the divergence of the heat transferred (i.e. $\frac{\partial}{\partial t}u(x, y, z, t) = \kappa_2(\nabla \bullet \vec{h}(x, y, z, t))$ for a positive constant κ_2). Combining both relationships, we have $\frac{\partial}{\partial t}u(x, y, z, t) = \nabla \bullet (\nabla u(x, y, z, t)) = \kappa_1\kappa_2(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})u(x, y, z, t)$. Thus, for some constant $k = \kappa_1\kappa_2$ (that is of course positive), the differential equation that describes the change of the temperature across the body over time is:

$$\frac{\partial u}{\partial t} = k\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right).$$

Of course, we do assume that the heat transfer at the surface of the body is negligible, so there is little to no energy loss — we could also consider unrealistic infinitely large bodies. For the sake of simplicity, we will only be considering the one-dimensional version of the heat equation, $\frac{\partial u}{\partial t} = k\frac{\partial u}{\partial^2 x}$, which applies to any body in which temperature does not vary in the y and z directions. We also impose the boundary conditions that both $u(0, t)$ and $u(l, t)$ are 0 (or any constant determined by the fixed temperature of the surrounding environment), which models a thin rod whose ends are of equal temperature, and that $u(x, 0)$ is equal to a certain function $f(x)$, which represents the distribution of temperature at some initial state. Although f is technically only defined over $[0, l]$, it can be generalized to a periodic function over all real numbers such that $f(x) = f(x + l)$ (as $f(0) = f(l)$).

Ignore the $u(x,0) = f(x)$ condition at this point. Solving the heat equation when u is a separable function (i.e. when u can be expressed as $X(x)T(t)$) is standard procedure: The differential equation becomes $X(x)T'(t) = kX''(x)T(t)$, or $\frac{1}{k}\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}$. Since one side depends on t and the other on x, this equation can only be satisfied if $\frac{1}{k}\frac{T'(t)}{T(t)}$ and $\frac{X''(x)}{X(x)}$ equal a constant $-\ell$, which must be negative, as otherwise T and hence u diverge to ∞ over time. We then obtain the linear differential equations $T'(t)+k\ell T(t) = 0$ and $X''(x)+\ell X(x) = 0$, which have the general solutions $T(t) = T_0 e^{-k\ell t}$ and $X(x) = X_1 e^{i\sqrt{\ell}x} + X_2 e^{-i\sqrt{\ell}x} = X_1' \cos(\sqrt{\ell}x) + X_2' \sin(\sqrt{\ell}x)$. Thanks to the boundary conditions, the only possibilities for $X(x)$ are of the form $X_0 \sin(\frac{n\pi}{l}x)$ for integer $n \geq 0$, forcing ℓ to be $\frac{n^2\pi^2}{l^2}$. We obtain the following solutions to the heat equation for each possible n:

$$u_n(x,t) = X(x)T(x) = U_0 e^{-\frac{kn^2\pi^2}{l^2}t} \sin(\frac{n\pi}{l}x).$$

We can create linear combinations of these solutions to obtain the general and not necessarily separable solution $C_0 + \sum_{n=1}^{\infty} C_n e^{-\frac{kn^2\pi^2}{l^2}t} \sin(\frac{n\pi}{l}x)$, where some values of n may be redundant (e.g. if l is an integer) and so $C_n = 0$. This kind of expression should be familiar — indeed, it is a trigonometric polynomial and is a Fourier series, so we hope that the initial condition $u(x,0) = f(x)$ can fix the coefficients C_n. Since $u(x,0) = C_0 + \sum_{n=1}^{\infty} C_n \sin(\frac{n\pi}{l}x) = C_0 + \sum_{n=1}^{\infty} C_n \frac{e^{i\frac{n\pi}{l}x} - e^{-i\frac{n\pi}{l}x}}{2i}$ must be equal to $f(x)$ and since the solution has a period of l, we conclude that the C_n must equal $\frac{2i}{l}\int_0^l f(x)e^{-i\frac{n\pi}{l}x}\,dx$, the coefficients in the Fourier series for f. (The only exception is C_0, which is simply $\int_0^l f(x)\,dx$.) Of course, the Fourier series $u(x,0)$ only converges to the initial state $f(x)$ pointwise under the Dirichlet-Jordan criterion. If f is differentiable and integrable (which are reasonable assumptions for real-life temperature distributions), then the pointwise convergence is guaranteed. Fourier's initial 1822 derivation did not show that the general solution always converges, resulting in controversy among the foremost physicists at the time, including Lagrange, Laplace, and Poisson (the same physicist who proved the aforementioned summation formula).

As an example, consider the evolution of the following body, of which only one segment had been heated, over time.

We can take a similar approach to derive a set of solutions to the three-dimensional heat equation, if we again assume that the body has a temperature of 0 (or some other constant) over the entire boundary. However, we can not easily apply the Fourier series as the initial state function would have three variables, and if we generalize the series to such functions (through double or triple integrals of a similar form), we would need to prove that the appropriate convergence criteria hold.

6 Conclusion

We have seen that the Fourier series for a given function can decompose it into a summation of exponentials of different integral frequencies, and this sum can accurately approximate the original function as long as certain conditions (e.g. the Dirichlet-Jordan criteria) are satisfied. A continuous counterpart is the Fourier transform, which describes the coefficients for the required exponentials of any frequency. The summation is now replaced by a continuous integral, and the latter, according to the Fourier Inversion Theorem, is equal to the original function. Some implications of this result we have proven include the Fourier transform's unitarity or the Poisson Summation Formula. We also demonstrated the utility of Fourier series in solving partial differential equations, specifically the heat equation of thermodynamics, but unfortunately we are not able to cover all the indispensable applications of Fourier analysis throughout other fields: number theory (functional equation derivation for the Riemann zeta function), quantum mechanics (Schrödinger wave equation solutions), statistics (serial correlation of time series), and computer science (compression of information in images), among many others.

References

[1] J. R. Cannon. *The One-Dimensional Heat Equation*. Cambridge University Press, 1984.

[2] T. H. Fay and P. H. Kloppers. The Gibbs' phenomenon. *International Journal of Mathematical Education in Science and Technology*, 32:73–89, 2001.

[3] M. R. Murty. *Problems in Analytic Number Theory*. Springer, New York, New York, 2 edition, 2008.

[4] E.M. Stein and R. Shakarchi. *Fourier Analysis: An Introduction*. Princeton University Press, 2003.