

The intuition behind the Fourier and Laplace transforms

Peter Haggstrom
mathsatbondibeach@gmail.com
<https://gotohaggstrom.com>

May 1, 2020

1 Background to the Fourier transform

On the internet one frequently sees engineering, maths and physics students plaintively seeking some form of insightful explanation of the "intuition" behind the Fourier and Laplace transforms. It is not uncommon to see comments along the lines of "I'm an electrical engineer but no-one has ever explained where the transforms came from". The reality is that a subject like Fourier theory is simply so immense that in a one semester course there is not enough time to give an historical context to the subject. Nevertheless, there are some fundamental aspects of the Fourier and Laplace transforms which can be quickly communicated, however, the "intuition" will depend on the background of who is doing the explaining.

Thus you will get a different type of emphasis from a functional analyst than from someone at the coal face of radio astronomy. Compare for instance how an expert in harmonic analysis, Eli Stein, explains Fourier theory at undergraduate level in his Princeton series of textbooks to how an expert in radio astronomy, Ron Bracewell, approaches the subject in his influential book "The Fourier transform and its applications".

A significant part of the problem is the use of the word "intuition", which is a form of mathematical pretentiousness. There are vast slabs of mathematics where the "intuition" only really exists for experts in the field, but not for struggling undergraduates! Indeed, Fourier theory, when viewed in its historical context, is the classic example of this, since Fourier's original "intuition" was hotly contested by contemporaries of the stature of Lagrange and Poisson. Fourier's "intuition" was at once both insightful (as history has amply shown) and audacious: he boldly claimed that his approach was capable of "developing any function whatever in an infinite series of sines or cosines of multiple arcs" ([1], page 168). When Fourier published his theory in 1822 it certainly was not "intuitively" obvious that an arbitrary non-periodic function could be represented as a convergent infinite series of sines and cosines. Yet Fourier's original motivation is still completely valid: you can construct something complex from simpler building blocks.

After 200 hundred years of intense development of Fourier theory, there now exists a gigantic superstructure of mathematical and physical principles which can be used to explain Fourier theory in ways that simply did not exist in the 19th century. Even so, walking into a room and simply writing a Fourier transform on a board as though it appeared from the heavens, fully formed, will not satisfy any inquiring mind. But one has to start somewhere and if you have studied analysis and understand the Weierstrass approximation theorem (which dates from 1885 - when Weierstrass was 70 -ie long after Fourier published his heat work), I think you can build an "intuition" for the concept of a Fourier series approximating a function and then make the conceptual leap from an infinite series to a continuous Fourier integral or transform. Courant and Hilbert actually did that in their influential book on mathematical physics ([2], pages 65-82). They explicitly proved that every continuous function $f(x)$ on $[-\pi, \pi]$ and for which $f(-\pi) = f(\pi)$, may be approximated uniformly by trigonometric polynomials $\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$ where the a_k, b_k are constants. Since both sines and cosines can be represented as power series ie infinite polynomial sums, the Weierstrass approximation route does make sense in that context.

Courant and Hilbert go on to take a function $f(x)$ represented by a Fourier series in the interval $-l < x < l$ as follows:

$$f(x) = \sum_{k=-\infty}^{\infty} a_k e^{ik\frac{\pi}{l}x} \quad (1)$$

where

$$a_k = \frac{1}{2l} \int_{-l}^l f(t) e^{-ik\frac{\pi}{l}x} dt \quad (2)$$

and they say "it seems desirable to let l go to ∞ , since then it is no longer necessary to require that f be continued periodically; thus one may hope to obtain a representation for a non-periodic function defined for all real x ([2], page 77).

On this basis it is proved in every Fourier theory course that $f(x)$ can be represented as follows:

$$f(x) = \frac{1}{\pi} \int_0^{\infty} du \int_{-\infty}^{\infty} f(t) \cos u(t-x) dt \quad (3)$$

So here is the high level intuition which is actually rigorous (subject to various caveats):

$f \rightarrow \text{Weierstrass approximation} = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \sin kx + b_k \cos kx) \rightarrow \text{Fourier integral}$

(4)

The coefficients are given by:

$$\begin{aligned}
 a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx \quad k = 1, 2, \dots \\
 b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx \quad k = 1, 2, \dots \\
 a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \, dx
 \end{aligned} \tag{5}$$

Even the use of the word "transform" is tricky since it suggests that you are changing or transforming something into something else and the next question is: "What do you do with the transform?". Well, if you can invert the transform you can get back to the original function. But why would you transform a function only to invert the result to get back to where you started? The quick and simple answer to this lay in the study of differential equations which is what drove Fourier and physics since Newton.

The Fourier transform of a derivative gives rise to multiplication in the transform space and the Fourier transform of a convolution integral gives rise to the product of Fourier transforms. The Fourier inversion theorem allows us to extract the original function. Such properties are extremely useful at a practical and theoretical level.

Here are the basic Fourier transformation rules, assuming f has suitable characteristics where $h \in \mathbb{R}$:

$$\begin{aligned}
 f(x+h) &\rightarrow \hat{f}(\xi)e^{2\pi ih\xi} \\
 f(x)e^{-2\pi ixh} &\rightarrow \hat{f}(\xi+h) \\
 f(\delta x) &\rightarrow \delta^{-1}\hat{f}(\delta^{-1}\xi) \text{ whenever } \delta > 0 \\
 f'(x) &\rightarrow 2\pi i\xi\hat{f}(\xi) \\
 -2\pi ix f(x) &\rightarrow \frac{\partial}{\partial \xi}\hat{f}(\xi) \\
 \widehat{f * g} &\rightarrow \hat{f}(\xi)\hat{g}(\xi)
 \end{aligned} \tag{6}$$

For the record I am using $\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ix\xi} \, dx$ as the Fourier transform.

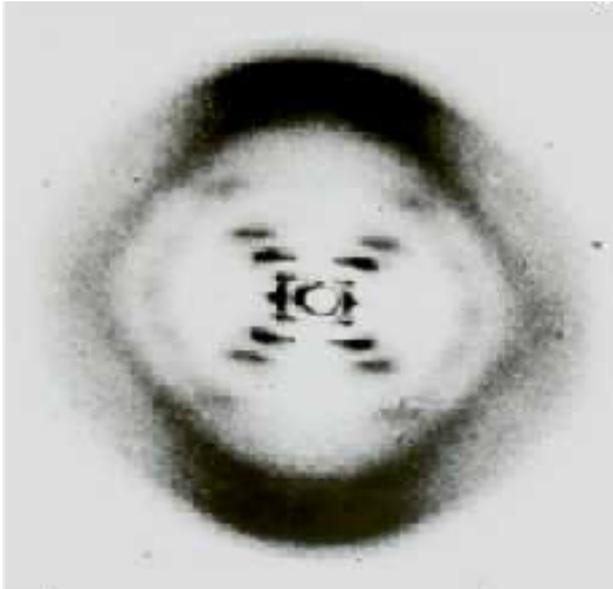
The heat equation has the form:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{7}$$

where $u = u(x, t)$ and we can take Fourier transforms of both sides of (7), then we can use the convolution property and then finally perform an inversion to get our solution. Some years ago I did a very detailed exposition of this process in [3]. It takes you through

every single step. A more recent paper seeks to motivate the Fourier transform through an analysis of the heat equation cases of discrete and continuous eigenvalues (see [5]).

Today the "intuition" in the world of spectroscopy, CT scanning and so on is that when you send an x-ray or some other sensing mechanism through something you get a "spectrum". The classic example is the famous DNA x-ray image obtained by Rosalind Franklin. The 2-dimensional "X" immediately suggested a double helix in 3 dimensions. And, yes, Fourier theory (and Bessel functions due to the cylindrical symmetry) played a pivotal role in that analysis.



Hence in the signal processing world it is quite "intuitive" to refer to transforms since the brain, bone etc do actually transform the signal which has to be reconstructed from its observed characteristics. This is why mathematicians work on inverse theory. The vast historical background now makes this mindset "intuitive".

It is now "intuitive" for electrical engineering students to take Fourier transforms of sines or cosines without realising how preposterous that is without distribution theory. Just so you understand this throwaway line, what is $\int_{-\infty}^{\infty} \cos x e^{-2\pi i x \xi} dx$? It clearly does not converge and when you plug that integral into Mathematica it confirms this. But when you use the command "FourierTransform" you miraculously get $\sqrt{\frac{\pi}{2}} \text{DiracDelta}[-1 + \omega] + \sqrt{\frac{\pi}{2}} \text{DiracDelta}[1 + \omega]$. This does not mean Fourier theory is flawed, rather it simply demonstrates its subtle nature and in 200 years it has reached a pedagogical standard whereby students can take on faith (and most do) that the Fourier transform of $\cos x$ actually makes sense. Some undergraduate courses actually have a go at showing why it is not all hocus pocus. It is applied mathematicians such as Ingrid Debauchies, *Stéphane* Mallat and Yves Meyer who have pushed this whole intellectual trend to the limits with wavelet theory.

For me the essence of a Fourier transform is that of approximation since this is the core concept behind real life examples such as transmitting a telephone call via an iPhone or Android phone. When you speak into the microphone there is a pressure change which affects the microphone's capacitance which in turn causes voltage to vary. The acoustic signal is thus a time varying voltage signal which is sampled discretely around 48,000 times per second with a 16 bit analog to digital converter circuit to provide a digital approximation of the analog voice signal (there is a "grid" and the samples are dumped into bins). The discrete Fourier transform plays a role in representing that signal in a digital form that can be efficiently transmitted with minimal error and then ultimately converted back (ie reconstructed) to an analog signal that you hear. There is a VAST theoretical and practical literature stretching back to Claude Shannon's work in 1948. The signal processing "mafia" will say something like this: "The Fourier transform is a method of expressing a given function of time (or any other appropriate co-ordinate, for that matter) in terms of a continuous set of exponential components of frequency. The resulting spectral-density function gives the relative weighting of each frequency component" ([4], page 92). The component functions used as building blocks of the approximation can be orthonormal sets of exponentials or things as pedestrian as step functions (think of the square step functions of Haar wavelets as an example).

Of course, once you have done a course in Fourier theory it is "intuitive" to represent a function by means of some sort of orthonormal basis but that is only because you have learnt all the foundations of the "intuition" that is embedded in 200 years of Fourier theory and its applications.

2 The intuition behind the Laplace transform

The logic behind the Laplace transform takes a different route compared to the logic which underpins the Fourier transform. Laplace developed a theory of generating functions in his famous treatise on probability theory ("Analytic Theory of Probability") where he expressed a function u as follows:

$$u = a_0 + a_1t + a_2t^2 + \dots \quad (8)$$

He developed this theory in the study of finite difference equations and hence their solution. The focus in the generating function approach is the coefficients - the powers of t are in a sense mere placeholders. A unit "impulse" at time $t = 1$, say, which is zero everywhere else will pick out the coefficient a_1 , and in electrical engineering courses this approach is frequently followed. Conceptually, an infinite series such as (8) is a linear combination of the complete polynomial set of functions $\{1, t, t^2, t^3, \dots\}$ so a function $f(t)$ can be approximated arbitrarily close in the mean square sense that $\int (f - \sum_{k=1}^n a_k t^k)^2 dt$ is less than any given $\epsilon > 0$. In this sense the generating function approach shares some of the motivation that was behind Fourier's original thinking. He was well aware of Laplace's work. In an electrical engineering course students are presented with a Laplace transform as a means of transforming the solution to a differential

equation into algebra.

So imagine now that we have a power series representation of some function ie $A(x) = \sum_{k=0}^{\infty} a(k)x^k$ where the $a(k)$ are constants. These constants can be viewed as the values of a discrete function initially and then the leap to the continuous domain can be made. For instance, if $a(k) = 1$ for all $k = 0, 1, 2, \dots$ and $|x| < 1$ we will have a convergent power series:

$$A(x) = 1 + x + x^2 + x^3 + \dots = \frac{1}{1-x} \quad (9)$$

Similarly if $a(k) = \frac{1}{k!}$ we will have:

$$A(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = e^x \quad (10)$$

and this holds for all real x .

We can recast the standard power series representation into a form which involves powers of x as follows:

$$\begin{aligned} A(x) &= \sum_{k=0}^{\infty} a(k)x^k \\ &= \sum_{k=0}^{\infty} a(k)e^{k \ln x} \\ &= \sum_{k=0}^{\infty} a(k)e^{-sk} \\ &= \sum_{k=0}^{\infty} a(k)e^{-sk} \underbrace{\Delta k}_{=1} \end{aligned} \quad (11)$$

All that has been done here is we have used the fact that $x = e^{\ln x}$ and since we want the infinite sum to converge we need $0 < x < 1$ so that $\ln x < 0$ and hence $s = -\ln x > 0$ or $-s = \ln x < 0$. Note that in this development if $x < 0$ and k takes on a value such as $\frac{1}{2}$ you would get a complex number, and for the purpose of developing this "intuition" we are staying away from that possibility! As in the case of the Fourier transform we imagine the variable k becoming continuous so the infinite sum morphs into an integral and we can view the coefficients as the value of a function and write $f(t)$, say, in the continuous limit. Equation (11) can be viewed as an approximation to an integral with rectangles having a base of $\Delta k = 1$ since the components are valued at $k = 0, 1, 2, \dots$. Thus we can write (11) suggestively as $\sum_{k=0}^{\infty} a(k)e^{-sk} \times \Delta k$, and we can then see that this sum becomes an integral transform of $f(t)$ as we decrease the rectangle width as k becomes smaller and takes on continuous values so that $a(n) \rightarrow f(t)$:

$$\mathcal{L}f(s) = \int_0^{\infty} f(t) e^{-st} dt \quad (12)$$

The theory of Laplace transforms has come a long way from this modest beginning since the properties have been heavily generalized and when you look at Laplace's original works you will not find the sort of development or intellectual progression that is standard in today's undergraduate engineering and maths courses.

However, we can see how Laplace's generating function approach is the driving force behind the modern approaches to differential equations by considering a basic difference equation. The aim is to find a sequence u_n such that $u_0 = 1$ and for all $n \geq 0$:

$$(n+1)u_{n+1} + u_n = 0 \quad (13)$$

The generating function is written as follows to emphasise the linkage to the coefficients:

$$(\mathcal{G}\mathbf{a})(t) = \sum_{n=0}^{\infty} a_n t^n \quad (14)$$

For every t we have from (13):

$$(n+1)u_{n+1}t^n + u_n t^n = 0 \quad (15)$$

In this problem we can write:

$$(\mathcal{G}\mathbf{u})(t) = u_0 + \sum_{n=0}^{\infty} u_{n+1}t^{n+1} = u_0 + u_1t + u_2t^2 + u_3t^3 + \dots \quad (16)$$

Now in the spirit of gaining an intuition we can send the Rigour Police on a long boozy holiday and differentiate the infinite series (16) with impunity:

$$(\mathcal{G}'\mathbf{u})(t) = u_1 + 2u_2t + 3u_3t^2 + \dots = \sum_{n=0}^{\infty} (n+1)u_{n+1}t^n \quad (17)$$

With this (15) becomes:

$$(\mathcal{G}'\mathbf{u})(t) + (\mathcal{G}\mathbf{u})(t) = 0 \quad (18)$$

Equation (18) is a garden variety ordinary differential equation whose solution is:

$$(\mathcal{G}\mathbf{u})(t) = Ae^{-t} \quad (19)$$

where A is a constant. With $t = 0$ we have that $1 = u_0 = (\mathcal{G}\mathbf{u})(0) = Ae^0 = A$ hence:

$$(\mathcal{G}\mathbf{u})(t) = e^{-t} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^n \quad (20)$$

Hence we have $u_n = \frac{(-1)^n}{n!}$.

3 References

[1] Joseph Fourier, The Analytical Theory of Heat, translated by Alexander Freeman, Cambridge University Press, 1878. The translation can be accessed here: <https://archive.org/details/analyticaltheory00fourrich>

[2] R Courant and D Hilbert, Methods of Mathematical Physics, Volume 1, Wiley, 1989.

[3] Peter Haggstrom, Basic Fourier Integrals ,<https://gotohaggstrom.com/Basic%20Fourier%20integrals.pdf>

[4] Ferrel G. Stremler, Introduction to Communication Systems, Second Edition, Wiley, 1982.

[5] Peter Haggstrom, , Using the heat equation to motivate the idea of the Fourier transform , <https://gotohaggstrom.com/Using%20the%20heat%20equation%20to%20motivate%20the%20idea%20of%20the%20Fourier%20transform.pdf>

4 History

Created 09 April 2020

22 April 2020 - added missing power of 2 in this $\int (f - \sum_{k=1}^n a_k t^k)^2 dt$

01 May 2020 - added reference to paper [5]