1

Physics and Fourier transforms

1.1 The qualitative approach

Ninety percent of all physics is concerned with vibrations and waves of one sort or another. The same basic thread runs through most branches of physical science, from acoustics through engineering, fluid mechanics, optics, electromagnetic theory and X-rays to quantum mechanics and information theory. It is closely bound to the idea of a *signal* and its *spectrum*. To take a simple example: imagine an experiment in which a musician plays a steady note on a trumpet or a violin, and a microphone produces a voltage proportional to the instantaneous air pressure. An oscilloscope will display a graph of pressure against time, $F(t)$, which is periodic. The reciprocal of the period is the frequency of the note, 440 Hz, say, for a well-tempered middle A – the tuning-up frequency for an orchestra.

The waveform is not a pure sinusoid, and it would be boring and colourless if it were. It contains 'harmonics' or 'overtones': multiples of the fundamental frequency, with various amplitudes and in various phases,[1] depending on the timbre of the note, the type of instrument being played and on the player. The waveform can be *analysed* to find the amplitudes of the overtones, and a list can be made of the amplitudes and phases of the sinusoids which it comprises. Alternatively a graph, $A(\nu)$, can be plotted (the sound-spectrum) of the amplitudes against frequency (Fig. 1.1).

$A(\nu)$ **is the Fourier transform of** $F(t)$.

Actually it is the *modular* transform, but at this stage that is a detail.

Suppose that the sound is not periodic – a squawk, a drumbeat or a crash instead of a pure note. Then to describe it requires not just a set of overtones

[1] 'Phase' here is an angle, used to define the 'retardation' of one wave or vibration with respect to another. One wavelength retardation, for example, is equivalent to a phase difference of 2π. Each harmonic will have its own phase, ϕ_m, indicating its position within the period.

Cambridge University Press
978-0-521-17683-5 - A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering, Third Edition
J. F. James
Excerpt
More information

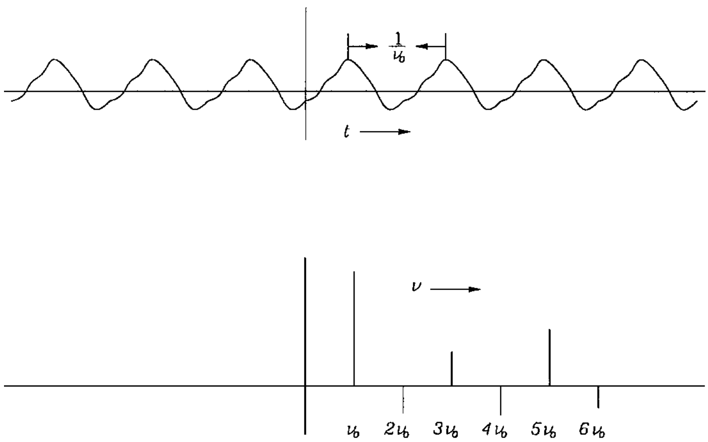

Fig. 1.1. The spectrum of a steady note: fundamental and overtones.

with their amplitudes, but a continuous range of frequencies, each present in an infinitesimal amount. The two curves would then look like Fig. 1.2.

The uses of a Fourier transform can be imagined: the identification of a valuable violin; the analysis of the sound of an aero-engine to detect a faulty gear-wheel; of an electrocardiogram to detect a heart defect; of the light curve of a periodic variable star to determine the underlying physical causes of the variation: all these are current applications of Fourier transforms.

1.2 Fourier series

For a steady note the description requires only the fundamental frequency, its amplitude and the amplitudes of its harmonics. A discrete sum is sufficient. We could write

$$F(t) = a_0 + a_1 \cos(2\pi\nu_0 t) + b_1 \sin(2\pi\nu_0 t) + a_2 \cos(4\pi\nu_0 t) + b_2 \sin(4\pi\nu_0 t) + a_3 \cos(6\pi\nu_0 t) + \cdots ,$$

where ν_0 is the fundamental frequency of the note. Sines as well as cosines are required because the harmonics are not necessarily 'in step' (i.e. 'in phase') with the fundamental or with each other.

More formally:

$$F(t) = \sum_{n=-\infty}^{\infty} a_n \cos(2\pi n\nu_0 t) + b_n \sin(2\pi n\nu_0 t) \tag{1.1}$$

and the sum is taken from $-\infty$ to ∞ for the sake of mathematical symmetry.

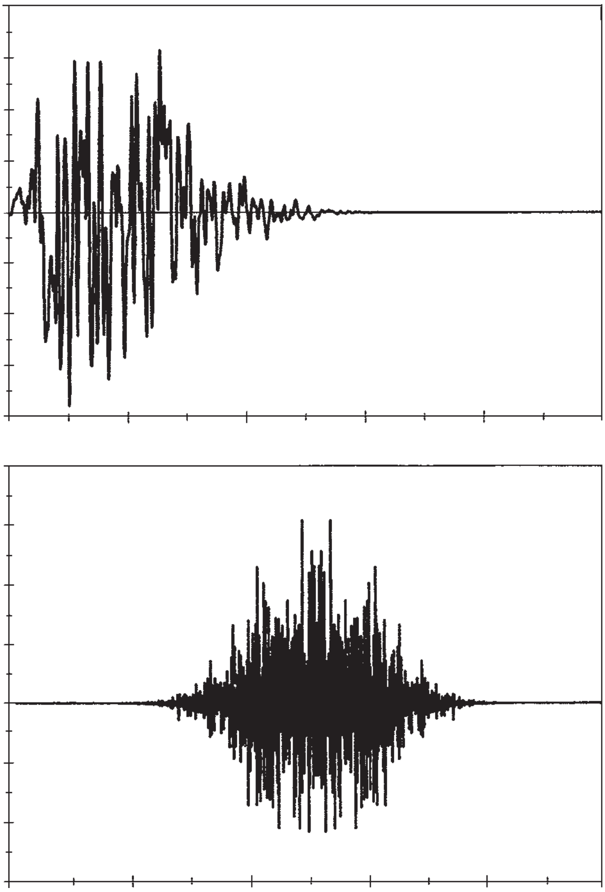

Fig. 1.2. The spectrum of a crash: all frequencies are present.

This process of constructing a waveform by adding together a fundamental frequency and overtones or harmonics of various amplitudes is called Fourier synthesis.

There are alternative ways of writing this expression: since $\cos x = \cos(-x)$ and $\sin x = -\sin(-x)$ we can write

$$F(t) = A_0/2 + \sum_{n=1}^{\infty} A_n \cos(2\pi n\nu_0 t) + B_n \sin(2\pi n\nu_0 t) \tag{1.2}$$

and the two expressions are identical, provided that we set $A_n = a_{-n} + a_n$ and $B_n = b_n - b_{-n}$. A_0 is divided by two to avoid counting it twice: as it is, A_0 can be found by the same formula that will be used to find all the A_n's.

Mathematicians and some theoretical physicists write the expression as

$$F(t) = A_0/2 + \sum_{n=1}^{\infty} A_n \cos(n\omega_0 t) + B_n \sin(n\omega_0 t)$$

and there are entirely practical reasons, which are discussed later, for *not* writing it this way.

1.3 The amplitudes of the harmonics

The alternative process – of extracting from the signal the various frequencies and amplitudes that are present – is called *Fourier analysis* and is much more important in its practical physical applications. In physics, we usually find the curve $F(t)$ experimentally and we want to know the values of the amplitudes A_m and B_m for as many values of m as necessary. To find the values of these amplitudes, we use the *orthogonality* property of sines and cosines. This property is that, if you take a sine and a cosine, or two sines or two cosines, each a multiple of some fundamental frequency, multiply them together and integrate the product over one period of that frequency, the result is always zero except in special cases.

If $P = 1/\nu_0$ is one period, then

$$\int_{t=0}^{P} \cos(2\pi n\nu_0 t) \cdot \cos(2\pi m\nu_0 t) dt = 0$$

and

$$\int_{t=0}^{P} \sin(2\pi n\nu_0 t) \cdot \sin(2\pi m\nu_0 t) dt = 0$$

unless $m = \pm n$, and

$$\int_{t=0}^{P} \sin(2\pi n\nu_0 t) \cdot \cos(2\pi m\nu_0 t) dt = 0$$

always.

The first two integrals are both equal to $1/(2\nu_0)$ if $m = n$.

We multiply the expression (1.2) for $F(t)$ by $\sin(2\pi m\nu_0 t)$ and the product is integrated over one period, P:

$$\int_{t=0}^{P} F(t)\sin(2\pi m\nu_0 t) dt = \frac{A_0}{2}\int_{t=0}^{P} \sin(2\pi m\nu_0 t) dt$$
$$+ \int_{t=0}^{P} \sum_{n=1}^{\infty} \{A_n \cos(2\pi n\nu_0 t) + B_n \sin(2\pi n\nu_0 t)\}\sin(2\pi m\nu_0 t) dt \quad (1.3)$$

and all the terms of the sum vanish on integration except

$$\int_0^P B_m \sin^2(2\pi m\nu_0 t)dt = B_m \int_0^P \sin^2(2\pi m\nu_0 t)dt$$
$$= B_m/(2\nu_0) = B_m P/2$$

so that

$$B_m = (2/P)\int_0^P F(t)\sin(2\pi m\nu_0 t)dt \tag{1.4}$$

and, provided that $F(t)$ is known in the interval $0 \to P$, the coefficient B_m can be found. If an analytic expression for $F(t)$ is known, the integral can often be done. On the other hand, if $F(t)$ has been found experimentally, a computer is needed to do the integrations.

The corresponding formula for A_m is

$$A_m = (2/P)\int_0^P F(t)\cos(2\pi m\nu_0 t)dt. \tag{1.5}$$

The integral can start anywhere, not necessarily at $t = 0$, so long as it extends over one period.

Example: Suppose that $F(t)$ is a square-wave of period $1/\nu_0$, so that $F(t) = h$ for $t = -b/2 \to b/2$ and 0 during the rest of the period, as in Fig. 1.3. Then

$$A_m = 2\nu_0 \int_{-1/(2\nu_0)}^{1/(2\nu_0)} F(t)\cos(2\pi m\nu_0 t)dt$$
$$= 2h\nu_0 \int_{-b/2}^{b/2} \cos(2\pi m\nu_0 t)dt$$

and the new limits cover only that part of the cycle where $F(t)$ is different from zero.

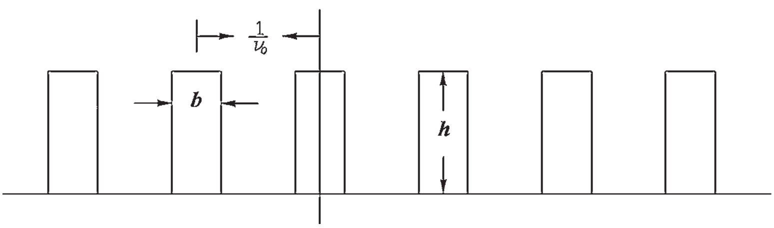

Fig. 1.3. A rectangular wave of period $1/\nu_0$ and pulse-width b.

If we integrate and put in the limits:

$$\begin{aligned} A_m &= \frac{2h\nu_0}{2\pi m\nu_0}\{\sin(\pi m\nu_0 b) - \sin(-\pi m\nu_0 b)\} \\ &= \frac{2h}{\pi m}\sin(\pi m\nu_0 b) \\ &= 2h\nu_0 b\{\sin(\pi\nu_0 mb)/(\pi\nu_0 mb)\}\,. \end{aligned}$$

All the B_n's are zero because of the symmetry of the function – we took the origin to be at the centre of one of the pulses.

The original function of time can be written

$$F(t) = h\nu_0 b + 2h\nu_0 b\sum_{m=1}^{\infty}\{\sin(\pi\nu_0 mb)/(\pi\nu_0 mb)\}\cos(2\pi m\nu_0 t) \tag{1.6}$$

or, alternatively,

$$F(t) = \frac{hb}{P} + \frac{2hb}{P}\sum_{m=1}^{\infty}\{\sin(\pi\nu_0 mb)/(\pi\nu_0 mb)\}\cos(2\pi m\nu_0 t). \tag{1.7}$$

Notice that the first term, $A_0/2$, is the *average* height of the function – the area under the top-hat divided by the period; and that the function $\sin(x)/x$, called 'sinc(x)', which will be described in detail later, has the value unity at $x = 0$, as can be shown using de l'Hôpital's rule.[2]

There are other ways of writing the Fourier series. It is convenient occasionally, though less often, to write $A_m = R_m \cos\phi_m$ and $B_m = R_m \sin\phi_m$, so that equation (1.2) becomes

$$F(t) = \frac{A_0}{2} + \sum_{m=1}^{\infty} R_m \cos(2\pi m\nu_0 t + \phi_m) \tag{1.8}$$

and R_m and ϕ_m are the amplitude and phase of the mth harmonic. A single sinusoid then replaces each sine and cosine, and the two quantities needed to define each harmonic are these amplitudes and phases in place of the previous A_m and B_m coefficients. In practice it is usually the amplitude, R_m, which is important, since the energy in an oscillator is proportional to the square of the amplitude of oscillation, and $|R_m|^2$ gives a measure of the power contained in each harmonic of a wave. 'Phase' is a simple and important idea. Two wave trains are 'in phase' if wave crests arrive at a certain point together. They are 'out of phase' if a trough from one arrives at the same time as the crest of the other. (Alternatively, they have 180° phase difference.) In Fig. 1.4 there are two

[2] De l'Hôpital's rule is that, if $f(x) \to 0$ as $x \to 0$ and $\phi(x) \to 0$ as $x \to 0$, the ratio $f(x)/\phi(x)$ is indeterminate, but is equal to the ratio $(df/dx)/(d\phi/dx)$ as $x \to 0$.

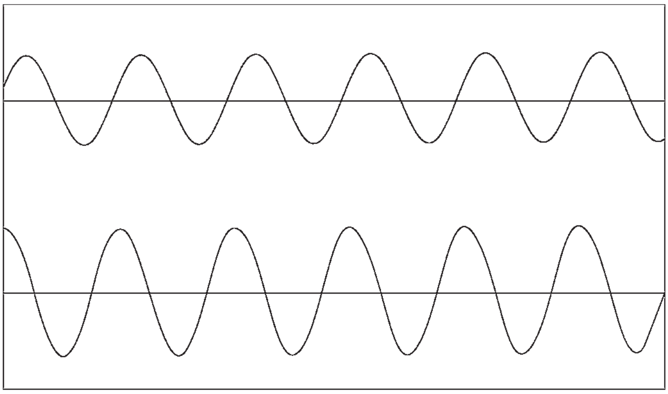

Fig. 1.4. Two wave trains with the same period but different amplitudes and phases. The upper has 0.7 times the amplitude of the lower and there is a phase-difference of 70°.

wave trains. The upper has 0.7 times the amplitude of the other and it *lags* (not *leads*, as it appears to do) the lower by 70°. This is because the horizontal axis of the graph is time, and the vertical axis measures the amplitude at a fixed point as it varies with time. Wave crests from the lower wave train arrive earlier than those from the upper. The important thing is that the 'phase-difference' between the two is 70°.

The most common way of writing the series expansion is with complex exponentials instead of trigonometrical functions. This is because the algebra of complex exponentials is easier to manipulate. The two ways are linked, of course, by de Moivre's theorem. We can write

$$F(t) = \sum_{-\infty}^{\infty} C_m e^{2\pi i m \nu_0 t},$$

where the coefficients C_m are now complex numbers in general and $C_m = C^*_{-m}$. (The exact relationship is given in detail in Appendix A.3.) The coefficients A_m, B_m and C_m are obtained from the *inversion formulae:*

$$A_m = 2\nu_0 \int_0^{1/\nu_0} F(t)\cos(2\pi m \nu_0 t)dt,$$

$$B_m = 2\nu_0 \int_0^{1/\nu_0} F(t)\sin(2\pi m \nu_0 t)dt,$$

$$C_m = 2\nu_0 \int_0^{1/\nu_0} F(t)e^{-2\pi m \nu_0 t}\, dt$$

(the minus sign in the exponent is important) or, if ω_0 has been used instead of ν_0 ($\nu_0 = \omega_0/(2\pi)$), then

$$A_m = (\omega_0/\pi)\int_0^{2\pi/\omega_0} F(t)\cos(m\omega_0 t)dt,$$

$$B_m = (\omega_0/\pi)\int_0^{2\pi/\omega_0} F(t)\sin(m\omega_0 t)dt,$$

$$C_m = (2\omega_0/\pi)\int_0^{2\pi/\omega_0} F(t)e^{-im\omega_0 t}\,dt.$$

The useful mnemonic form to remember for finding the coefficients in a Fourier series is

$$A_m = \frac{2}{\text{period}}\int_{\text{one period}} F(t)\cos\left\{\frac{2\pi mt}{\text{period}}\right\}dt, \tag{1.9}$$

$$B_m = \frac{2}{\text{period}}\int_{\text{one period}} F(t)\sin\left\{\frac{2\pi mt}{\text{period}}\right\}dt \tag{1.10}$$

and remember that the integral can be taken from any starting point, a, provided that it extends over one period to an upper limit $a + P$. The integral can be split into as many subdivisions as needed if, for example, $F(t)$ has different analytic forms in different parts of the period.

1.4 Fourier transforms

Whether $F(t)$ is periodic or not, a complete description of $F(t)$ can be given using sines and cosines. If $F(t)$ is not periodic it requires all frequencies to be present if it is to be synthesized. A non-periodic function may be thought of as a limiting case of a periodic one, where the period tends to infinity, and consequently the fundamental frequency tends to zero. The harmonics are more and more closely spaced and in the limit there is a continuum of harmonics, each one of infinitesimal amplitude, $a(\nu)d\nu$, for example. The summation sign is replaced by an integral sign and we find that

$$F(t) = \int_{-\infty}^{\infty} a(\nu)d\nu\cos(2\pi\nu t) + \int_{-\infty}^{\infty} b(\nu)d\nu\sin(2\pi\nu t) \tag{1.11}$$

or, equivalently,

$$F(t) = \int_{-\infty}^{\infty} r(\nu)\cos(2\pi\nu t + \phi(\nu))d\nu \tag{1.12}$$

or, again,

$$F(t) = \int_{-\infty}^{\infty} \Phi(\nu)e^{2\pi i\nu t}\,d\nu. \tag{1.13}$$

If $F(t)$ is real, that is to say, if the insertion of any value of t into $F(t)$ yields a real number, then $a(\nu)$ and $b(\nu)$ are real too. However, $\Phi(\nu)$ may be complex and indeed will be if $F(t)$ is asymmetrical so that $F(t) \neq F(-t)$. This can sometimes cause complications, and these are dealt with in Chapter 8: but $F(t)$ is often symmetrical and then $\Phi(\nu)$ is real and $F(t)$ comprises only cosines. We *could* then write

$$F(t) = \int_{-\infty}^{\infty} \Phi(\nu)\cos(2\pi\nu t)d\nu$$

but, because complex exponentials are easier to manipulate, we take equation (1.13) above as the standard form. Nevertheless, for many practical purposes only real and symmetrical functions $F(t)$ and $\Phi(\nu)$ need be considered.

Just as with Fourier series, the function $\Phi(\nu)$ can be recovered from $F(t)$ by inversion. This is the cornerstone of Fourier theory because, astonishingly, the inversion has exactly the same form as the synthesis, and we can write, if $\Phi(\nu)$ is real and $F(t)$ is symmetrical,

$$\Phi(\nu) = \int_{-\infty}^{\infty} F(t)\cos(2\pi\nu t)dt, \tag{1.14}$$

so that not only is $\Phi(\nu)$ the Fourier transform of $F(t)$, but also $F(t)$ is the Fourier transform of $\Phi(\nu)$. The two together are called a 'Fourier pair'.

The complete and rigorous proof of this is long and tedious[3] and it is not necessary here; but the formal definition can be given and this is a suitable place to abandon, for the moment, the physical variables time and frequency and to change to the pair of abstract variables, x and p, which are usually used. The formal statement of a Fourier transform is then

$$\Phi(p) = \int_{-\infty}^{\infty} F(x)e^{2\pi ipx}\,dx, \tag{1.15}$$

$$F(x) = \int_{-\infty}^{\infty} \Phi(p)e^{-2\pi ipx}\,dp \tag{1.16}$$

and this pair of formulae[4] will be used from here on.

[3] It is to be found, for example, in E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Clarendon Press, Oxford, 1962 or in R. R. Goldberg, *Fourier Transforms*, Cambridge University Press, Cambridge, 1965.

[4] Sometimes one finds

$$\Phi(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(x)e^{ipx}\,dx; \qquad F(x) = \int_{-\infty}^{\infty} \Phi(p)e^{-ipx}\,dp$$

as the defining equations, and again symmetry is preserved by some people by defining the transform by

$$\Phi(p) = \left\{\frac{1}{2\pi}\right\}^{1/2}\int_{-\infty}^{\infty} F(x)e^{ipx}\,dx; \qquad F(x) = \left\{\frac{1}{2\pi}\right\}^{1/2}\int_{-\infty}^{\infty} \Phi(p)e^{-ipx}\,dp.$$

Symbolically we write

$$\Phi(p) \rightleftharpoons F(x).$$

One and only one of the integrals must have a minus sign in the exponent. Which of the two you choose does not matter, so long as you keep to the rule. If the rule is broken half way through a long calculation the result is chaos; but if someone else has used the opposite choice, the Fourier pair calculated of a given function will be the complex conjugate of that given by your choice.

When time and frequency are the conjugate variables we shall use

$$\Phi(\nu) = \int_{-\infty}^{\infty} F(t)e^{-2\pi i\nu t}\, dt, \qquad (1.17)$$

$$F(t) = \int_{-\infty}^{\infty} \Phi(\nu)^{2\pi i\nu t}\, d\nu \qquad (1.18)$$

and again, symbolically,

$$\Phi(\nu) \rightleftharpoons F(t).$$

There are two good reasons for incorporating the 2π into the exponent. Firstly the defining equations are easily remembered without worrying where the 2π's go, but, more importantly, quantities like t and ν are actually physically measured quantities – time and frequency – rather than time and *angular* frequency, ω. Angular measure is for mathematicians. For example, when one has to integrate a function wrapped around a cylinder it is convenient to use the angle as the independent variable. Physicists will generally find it more convenient to use t and ν, for example, with the 2π in the exponent.

1.5 Conjugate variables

Traditionally x and p are used when abstract transforms are considered and they are called 'conjugate variables'. Different fields of physics and engineering use different pairs, such as frequency, ν, and time, t, in acoustics, telecommunications and radio; position, x, and momentum divided by Planck's constant, $p/\hbar$, in quantum mechanics; and aperture, x, and the sine of the diffraction angle divided by the wavelength, $p = \sin\theta/\lambda$, in diffraction theory.

In general we will use x and p as abstract entities and give them a physical meaning when an illustration seems called for. It is worth remembering that x and p have inverse dimensionality, as in time, t, and frequency, t^{-1}. The product px, like any exponent, is always a dimensionless number.

One further definition is needed: the 'power spectrum' of a function.[5] This notion is important in electrical engineering as well as in physics. If power

[5] Actually the *energy* spectrum; 'power spectrum' is just the conventional term used in most books. This is discussed in more detail in Chapter 4.