8 Fourier Transforms

Given a (sufficiently well behaved) function $f : S^1 \to \mathbb{C}$, or equivalently a periodic function on the real line, we've seen that we can represent f as a Fourier series

$$f(x) = \frac{1}{2L} \sum_{n\in\mathbb{Z}} \hat{f}_n \, \mathrm{e}^{\mathrm{i}n\pi x/L} \tag{8.1}$$

where the period is $2L$. In this chapter we'll extend these ideas to non-periodic functions $f : \mathbb{R} \to \mathbb{C}$. This extension, which is again due to Fourier, is one of the most important ideas in all Applied Mathematics (and a great deal of Pure). In one form or other it is at the heart of all spectroscopy (from crystallography to understanding the structure of proteins in your DNA), to the definition of a particle in Quantum Field Theory, to all image processing done by the Nvidia graphics chip in your iPad, not to mention Harmonic Analysis and a large chunk of Representation Theory.

To get started, recall that in the periodic case, the information about our function was stored in its list of Fourier coefficients $\hat{f}_n$, defined by

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}nx} \, f(x) \, \mathrm{d}x \tag{8.2}$$

for $n \in \mathbb{Z}$. For a (non-periodic) function $f : \mathbb{R} \to \mathbb{C}$, we likewise define the *Fourier transform* $\tilde{f}(k)$ of $f(x)$ to be[30]

$$\tilde{f}(k) := \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} \, f(x) \, \mathrm{d}x \, . \tag{8.3}$$

The Fourier transform is an example of a linear transform, producing an output function $\tilde{f}(k)$ from the input $f(x)$. We'll sometimes use the notation $\tilde{f} = \mathcal{F}[f]$, where the $\mathcal{F}$ on the *rhs* is to be viewed as the operation of 'taking the Fourier transform', *i.e.* performing the integral in (8.3).

8.1 Simple properties of Fourier transforms

The Fourier transform has a number of elementary properties. For any constants $c_1, c_2 \in \mathbb{C}$ and integrable functions f, g the Fourier transform is linear, obeying

$$\mathcal{F}[c_1 f + c_2 g] = c_1 \mathcal{F}[f] + c_2 \mathcal{F}[g] \, .$$

By changing variables in the integral, it is also readily verified that it obeys

$$\begin{aligned} \text{translation} \qquad & \mathcal{F}[f(x-a)] = \mathrm{e}^{-\mathrm{i}ka} \tilde{f}(k) \\ \text{re-phasing} \qquad & \mathcal{F}[\mathrm{e}^{\mathrm{i}\ell x} f(x)] = \tilde{f}(k-\ell) \\ \text{scaling} \qquad & \mathcal{F}[f(cx)] = \frac{1}{|c|} \tilde{f}(k/c) \, . \end{aligned}$$

[30]Be warned! Various authors use a factor of $1/2\pi$ or $1/\sqrt{2\pi}$ instead. The choice you make here is purely conventional, but affects the corresponding choice you make in the Fourier inversion theorem of section 8.2.

Furthermore, if we define the *convolution*[31] $f * g$ of two functions to be the integral

$$f * g(x) = \int_{-\infty}^{\infty} f(x-y)\, g(y)\, \mathrm{d}y \tag{8.4}$$

then, provided f and g are sufficiently well–behaved for us to change the order of integration, the Fourier transform is

$$\begin{aligned} \mathcal{F}[f * g(x)] &= \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} \left[\int_{-\infty}^{\infty} f(x-y)\, g(y)\, \mathrm{d}y \right] \mathrm{d}x = \int_{\mathbb{R}^2} \mathrm{e}^{-\mathrm{i}k(x-y)}\, f(x-y)\, \mathrm{e}^{-\mathrm{i}ky}\, g(y)\, \mathrm{d}x\, \mathrm{d}y \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}ku}\, f(u)\, \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}ky}\, g(y)\, \mathrm{d}y = \mathcal{F}[f]\, \mathcal{F}[g]\,. \end{aligned} \tag{8.5}$$

In other words, the Fourier transform of a convolution of two functions is the product of their Fourier transforms.

By far the most useful property of the Fourier transform comes from the fact that the Fourier transform 'turns differentiation into multiplication'. Specifically, the Fourier transform of the derivative f' of a (smooth, integrable) function f is given by

$$\mathcal{F}[f'(x)] = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\, f'(x)\, \mathrm{d}x = -\int_{-\infty}^{\infty} \left(\frac{d}{dx} \mathrm{e}^{-\mathrm{i}kx} \right) f(x)\, \mathrm{d}x = \mathrm{i}k\tilde{f}(k) \tag{8.6}$$

where we note that the assumption that $f(x)$ was integrable means that in particular it must decay as $|x| \to \infty$, so there is no boundary term. Notice also that, as a sort of converse, if $\tilde{f}(k) = \mathcal{F}[f(x)]$ then the Fourier transform of $xf(x)$ is given by

$$\mathcal{F}[xf(x)] = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\, xf(x)\, \mathrm{d}x = \mathrm{i}\frac{d}{dk} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\, f(x)\, \mathrm{d}x = \mathrm{i}\tilde{f}'(k) \tag{8.7}$$

provided $xf(x)$ is itself integrable (and thus $\tilde{f}$ is differentiable) so that we can justify differentiating under the integral sign.

The fact that differentiation *wrt* x becomes multiplication by $\mathrm{i}k$ is important because it allows us to take the Fourier transform of a differential equation. Suppose

$$\mathcal{L}(\partial) = \sum_{r=0}^{p} c_r \frac{d^r}{dx^r}$$

is a differential operator with constant coefficients $c_r \in \mathbb{C}$. Then if $y : \mathbb{R} \to \mathbb{C}$ has Fourier transform $\tilde{y}$, the Fourier transform of $\mathcal{L}y$ is

$$\mathcal{F}[\mathcal{L}(\partial)y] = \mathcal{L}(\mathrm{i}k)\tilde{y}(k) \tag{8.8}$$

where the differential operator $\mathcal{L}(\partial)$ has been replaced by muliplication by the polynomial

$$\mathcal{L}(\mathrm{i}k) = \sum_{r=0}^{p} c_r (\mathrm{i}k)^r\,.$$

[31] Convolution is 'Faltung' in German, which picturesquely describes the way in which the functions are combined — the graph of g is ßipped (or folded) about the variable vertical line $u = x/2$ and then integrated against f.

(We've also assumed that all the Fourier integrals we meet converge). Thus, if y obeys some differential equation $\mathcal{L}(\partial)y(x) = f(x)$ on $\mathbb{R}$, then we have simply

$$\tilde{y}(k) = \tilde{f}(k)/\mathcal{L}(\mathrm{i}k) \tag{8.9}$$

in terms of the Fourier transforms.

The power of this approach becomes most apparent in higher dimensions. For example, consider the pde

$$\nabla^2\phi(\mathbf{x}) - m^2\phi(\mathbf{x}) = -\rho(\mathbf{x}) \tag{8.10}$$

for $\phi : \mathbb{R}^n \to \mathbb{C}$, where ∇^2 is the n-dimensional Laplacian. Defining the Fourier transform for (appropriately integrable) functions on $\mathbb{R}^n$ by the obvious generalization

$$\mathcal{F}[\phi(\mathbf{x})] = \tilde{\phi}(\mathbf{k}) = \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\phi(\mathbf{x})\,\mathrm{d}^n x\,, \tag{8.11}$$

our pde becomes simply

$$\tilde{\phi}(\mathbf{k}) = \frac{\tilde{\rho}(\mathbf{k})}{|\mathbf{k}|^2 + m^2} \tag{8.12}$$

in terms of Fourier transforms.

8.2 The Fourier inversion theorem

We've just seen that linear differential equations can often become essentially trivial after a Fourier transform. However, if we are to make use of this result we need to be able to reconstruct the original function from knowledge of its Fourier transform. Recall that in the periodic case, we were able to represent

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{f}_n\,\mathrm{e}^{2\pi\mathrm{i}nx/L} \tag{8.13}$$

for sufficiently smooth functions (or more generally, distributions) $f : S^1 \to \mathbb{C}$ with period L. To extend this to the non–periodic case, let's imagine trying to taking a limit of (8.13) where the period $L \to \infty$. We define $\Delta k \equiv 2\pi/L$ and note that the coefficients in (8.13) are

$$\hat{f}_n = \frac{1}{L}\int_{-L/2}^{L/2} \mathrm{e}^{-\mathrm{i}n\Delta ky}\,f(y)\,\mathrm{d}y = \frac{\Delta k}{2\pi}\int_{-L/2}^{L/2} \mathrm{e}^{-\mathrm{i}n\Delta ky}\,f(y)\,\mathrm{d}y \tag{8.14}$$

For this integral to exist in the limit $L \to \infty$ we should require that $\int_{\mathbb{R}} |f(y)|$ exists, but in that case the $1/L$ factor implies that for each n, $\hat{f}_n \to 0$ too. Nevertheless, for any finite L we can substitute (8.14) into (8.13) to obtain

$$f(x) = \sum_{n\in\mathbb{Z}} \left(\frac{\Delta k}{2\pi}\int_{-L/2}^{L/2} \mathrm{e}^{\mathrm{i}n\Delta k(x-y)}\,f(y)\,\mathrm{d}y\right) \tag{8.15}$$

Now recall the Riemann sum definition of the integral of a function g: As $\Delta k \to 0$ we have

$$\lim_{\Delta k\to 0}\sum_{n\in\mathbb{Z}} \Delta k\, g(n\Delta k) = \int_{\mathbb{R}} g(k)\,\mathrm{d}k \tag{8.16}$$

with k becoming a continuous variable. For (8.15) we take[32]

$$g(n\Delta k) = \frac{\mathrm{e}^{\mathrm{i}n\Delta kx}}{2\pi}\int_{-L/2}^{L/2} \mathrm{e}^{\mathrm{i}n\Delta ky}\, f(y)\,\mathrm{d}y \tag{8.17}$$

with y being viewed as a parameter. Thus, letting $L \to \infty$ we get

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\left[\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}ky}\, f(y)\,\mathrm{d}y\right]\mathrm{d}x = \frac{1}{2\pi}\int_{\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\,\tilde{f}(k)\,\mathrm{d}k\,. \tag{8.18}$$

This result is known as the *Fourier inversion theorem*; if the manipulations above can be made rigorous, then the function $f(x)$ itself can be expressed in terms of an integral of the Fourier transformed function $\tilde{f}(k)$. In line with our earlier notation, we sometimes write $f(x) = \mathcal{F}^{-1}[\tilde{f}(k)]$ where the *rhs* is just shorthand for the integral in (8.18).

Notice that the inverse Fourier transform looks almost identical to the Fourier transform itself — the only difference is the sign in the exponent and the factor of $1/2\pi$. In particular, replacing x by $-x$ in (8.18) we have the *duality* property

$$\tilde{f}(k) = \mathcal{F}[f(x)] \qquad \Leftrightarrow \qquad f(-x) = \frac{1}{2\pi}\mathcal{F}[\tilde{f}(k)]\,. \tag{8.19}$$

This observation is very useful: if we recognize some specific function g as being the Fourier transform of some function f, then we can immediately write down the Fourier transform of g itself in terms of f. For example, we saw in equation (8.5) that $\mathcal{F}[f*g(x)] = \tilde{f}(k)\,\tilde{g}(k)$. It now follows that

$$\mathcal{F}[f(x)\,g(x)] = \frac{1}{2\pi}\int \tilde{f}(k-\ell)\,\tilde{g}(\ell)\,\mathrm{d}\ell = \frac{1}{2\pi}\tilde{f}*\tilde{g}(k) \tag{8.20}$$

so that the Fourier transform of a product of two functions is the convolution of their individual Fourier transforms.

The Fourier inversion theorem allows us, in principle, to complete the problem of finding the solution to a linear differential equation. We saw in (8.9) that if $\mathcal{L}(\partial)y(x) = f(x)$ then $\tilde{y}(k) = \tilde{f}(k)/\mathcal{L}(\mathrm{i}k)$ where $\mathcal{L}(\mathrm{i}k)$ is a polynomial in k. Provided the *rhs* is integrable so that the Fourier inversion theorem holds, we have now

$$y(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\,\frac{\tilde{f}(k)}{\mathcal{L}(\mathrm{i}k)}\,\mathrm{d}k\,. \tag{8.21}$$

As an example, the pde $\nabla^2\phi - m^2\phi = -\rho$ on $\mathbb{R}^n$ is solved by

$$\phi(\mathbf{x}) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \frac{\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\tilde{\rho}(\mathbf{k})}{|\mathbf{k}|^2+m^2}\,\mathrm{d}^n k\,. \tag{8.22}$$

Furthermore, from the result (8.20) that the Fourier transform of a product of functions is the convolution of the Fourier transforms, we see that our result will involve a convolution of the forcing term $f(x)$ with the inverse Fourier transform of the rational function $1/\mathcal{L}(\mathrm{i}k)$.

[32] The function g in (8.17) itself changes as $\Delta k \sim 1/L \to 0$, because the integral in its definition is taken over $[-L/2, L/2]$. However, it has a well defined limit $g_{\Delta k}(n\Delta k) \to g_0(k) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}ky}\, f(y)\,\mathrm{d}y$ and so the limit of (8.15) is still as given in (8.18).

In practice, the final step of actually carrying out the integrals in the inverse Fourier transform can often be quite tricky. We'll look at a few simple examples where we can guess the answer in section 8.4, but the inverse Fourier transform is usually best done with the aid of techniques from Complex Analysis that you'll meet next term.

8.2.1 Parseval's theorem for Fourier transforms

A further important feature of the Fourier transform is that it preserves the (L^2) inner product of two functions – just as we saw for Fourier series. To see this, suppose $f, g : \mathbb{R}^n \to \mathbb{C}$ and sufficiently well–behaved that the Fourier transforms $\tilde{f}$ and $\tilde{g}$, exist and that they can themselves be represented in terms of $\tilde{f}$ and $\tilde{g}$ using the Fourier inversion theorem. Then

$$
\begin{aligned}
(f, g) &= \int_{\mathbb{R}^n} f^*(x)\, g(x)\, \mathrm{d}^n x = \int_{\mathbb{R}^n} f^*(x) \left[\frac{1}{2\pi}\int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x}\, \tilde{g}(k)\, \mathrm{d}^n k\right] \mathrm{d}^n x \\
&= \frac{1}{2\pi}\int_{\mathbb{R}^n} \tilde{g}(k) \left[\int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x}\, f^*(x)\, \mathrm{d}^n x\right] \mathrm{d}^n k \\
&= \frac{1}{2\pi}\int_{\mathbb{R}^n} \tilde{g}(k) \left[\int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot x}\, f(x)\, \mathrm{d}^n x\right]^* \mathrm{d}^n k = \frac{1}{2\pi}\int_{\mathbb{R}^n} \tilde{g}(k)\, \tilde{f}(k)^*\, \mathrm{d}^n k \\
&= \frac{1}{2\pi}(\tilde{f}, \tilde{g})\,,
\end{aligned}
\tag{8.23}
$$

where in going to the second line we changed the order of the integrals — this step is justified since we've assumed that f and $\tilde{g}$ are both absolutely integrable on $\mathbb{R}^n$ in order for the Fourier transforms and inverse Fourier transforms to exist.

As a special case, if $f = g$ then we obtain

$$
(f, f) = \frac{1}{2\pi}(\tilde{f}, \tilde{f}) \tag{8.24}
$$

so that the L^2-norm of a Fourier transform agrees with the L^2-norm the original function, up to a factor of $\sqrt{2\pi}$.[33] This is Parseval's theorem in the context of non–periodic functions.

8.3 The Fourier transform of Schwartz functions and tempered distributions

This section is again non–examinable, though you will need to know the results of subsection 8.3.1.

In this section we'll give a more careful discussion of when the Fourier transform exists, and when the manipulations in the above derivation of the Fourier inversion theorem can be made rigorous.

The first question we should ask is 'for what type of object does the Fourier transform make sense?'. To get started, suppose that $f : \mathbb{R} \to \mathbb{C}$ is Riemann integrable on every interval $[a, b]$ and that $\int_{-\infty}^{\infty} |f(x)|\, \mathrm{d}x$ exists. Then whenever k is real, we have

$$
|\tilde{f}(k)| = \left|\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\, f(x)\, \mathrm{d}x\right| \leq \int_{-\infty}^{\infty} |\mathrm{e}^{-\mathrm{i}kx}\, f(x)|\, \mathrm{d}x = \int_{-\infty}^{\infty} |f(x)|\, \mathrm{d}x \tag{8.25}
$$

[33] Some authors choose to include factors of $1/\sqrt{2\pi}$ in their definition of the Fourier transform to avoid this factor here.

which converges by our assumptions on f. In fact, it's straightforward to show that with these assumptions on f, $\tilde{f}(k)$ is in fact everywhere *continuous* (I recommend this as an excercise if you like Analysis). However, even if $\int_{\mathbb{R}} |f(x)|\,\mathrm{d}x$ exists, it can often be that $\int_{\mathbb{R}} |\tilde{f}(k)|\,\mathrm{d}k$ diverges. For example, suppose $f(x)$ is the 'top hat function'

$$f(x) = \begin{cases} 1 & |x| \leq 1 \\ 0 & \text{else}\,. \end{cases} \tag{8.26}$$

This is certainly integrable, and indeed we readily compute the Fourier transform to be

$$\tilde{f}(k) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\, f(x)\,\mathrm{d}x = \frac{1}{2\pi}\int_{-1}^{1} \mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x = 2\,\frac{\sin k}{k}\,. \tag{8.27}$$

The problem is that $\tilde{f}(k)$ is now not integrable, since

$$\begin{aligned} \int_{-(N+1)\pi}^{(N+1)\pi} |\tilde{f}(k)|\mathrm{d}k &\geq \sum_{n=0}^{N} \int_{(n+\frac{1}{4})\pi}^{(n+\frac{3}{4})\pi} |\tilde{f}(k)|\mathrm{d}k \geq \sum_{n=0}^{N} \int_{(n+\frac{1}{4})\pi}^{(n+\frac{3}{4})\pi} \frac{1}{\sqrt{2}k}\mathrm{d}k \\ &\geq \sum_{n=0}^{N} \frac{\pi/2}{\sqrt{2}(n+1)\pi} = \frac{1}{2\sqrt{2}}\sum_{n=0}^{N}\frac{1}{n+1} \end{aligned} \tag{8.28}$$

which diverges as $N \to \infty$. But if taking the Fourier transform of an integrable function $f(x)$ can lead to a function $\tilde{f}(k)$ that is not itself integrable, how do we make sense of the Fourier inversion theorem?

Recall from section 6.3 that a *Schwartz function* ψ is an infinitely smooth function obeying

$$\sup_{x\in\mathbb{R}} \; x^m\,\psi^{(n)}(x) < \infty \qquad \text{for all } m, n \in \mathbb{N}. \tag{8.29}$$

so that in particular ψ decay rapidly as $|x| \to \infty$. As before,

$$|\tilde{\psi}(k)| = \left|\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\psi(x)\,\mathrm{d}x\right| \leq \int_{-\infty}^{\infty} \left|\mathrm{e}^{-\mathrm{i}kx}\,\psi(x)\right|\mathrm{d}x = \int_{-\infty}^{\infty} |\psi(x)|\,\mathrm{d}x < \infty \tag{8.30}$$

and so that $\tilde{\psi}(k)$ is bounded for all $k \in \mathbb{R}$. What is more, since all the functions $x^m\psi^{(n)}(x)$ of involving monomials times derivatives of a Schwartz function are themselves infinitely smooth and bounded, their Fourier transforms also exist. Explicitly, we have

$$\tilde{\psi}^{(m)}(k) = \frac{d^m}{dk^m}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\psi(x)\,\mathrm{d}x = \mathcal{F}[(-\mathrm{i}x)^m\psi(x)] \tag{8.31}$$

where we are justified in differentiating under the integral since $x^m\psi(x)$ is indeed integrable as $\psi(x)$ is a Schwartz function. In particular, since $\mathcal{F}[x^m\psi(x)]$ is bounded for all $k \in \mathbb{R}$, all the derivatives of $\tilde{\psi}$ exist and are bounded, so $\tilde{\psi}$ is infinitely smooth. Likewise

$$k^n\tilde{\psi}(k) = \mathrm{i}^n \int_{-\infty}^{\infty} \frac{d^n\mathrm{e}^{-\mathrm{i}kx}}{dx^n}\,\psi(x)\,\mathrm{d}x = (-\mathrm{i})^n\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\psi^{(n)}(x)\,\mathrm{d}x = (-\mathrm{i})^n\mathcal{F}[\psi^{(n)}(x)] \tag{8.32}$$

which is again bounded for all $k \in \mathbb{R}$. Combining these results shows that the Fourier transform $\tilde{\psi}(k)$ of a Schwartz function $\psi(x)$ is again a Schwartz function. Thus, the Fourier transform is a map $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$, and by iterating the duality (8.19) four times we find the map

$$\frac{1}{(2\pi)^4}\,\mathcal{F}^4 : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R}) \tag{8.33}$$

is actually the identity: Up to a numerical factor, taking the (forwards) Fourier transform of a Schwartz function four times in succession just reproduces the Schwartz function.

To go further and define Fourier transforms for more general objects, less well–behaved than Schwartz functions, we first note that if $\phi, \psi \in \mathcal{S}(\mathbb{R})$ then the derivation of (8.23) is sound and

$$(\mathcal{F}\phi, \mathcal{F}\psi) = 2\pi(\phi, \psi)\,. \tag{8.34}$$

Since $\mathcal{F}\circ\mathcal{F}^{-1} = 1$ for Schwartz functions, we can equivalently write this result as $(\mathcal{F}\phi, \tilde{\psi}) = 2\pi(\phi, \mathcal{F}^{-1}\tilde{\psi})$, or just

$$(\mathcal{F}\phi, \chi) = 2\pi(\phi, \mathcal{F}^{-1}\chi)\,, \tag{8.35}$$

where χ is again some Schwartz function. Now we recall from chapter 6 that we can treat $\phi \in \mathcal{S}(\mathbb{R})$ as a special case of a tempered distribution $T_\phi \in \mathcal{S}'(\mathbb{R})$, which acts on a test function ψ via $T_\phi[\psi] = (\phi, \psi)$. In the case that $T \in \mathcal{S}'(\mathbb{R})$ is a more general tempered distribution, we similarly define its Fourier transform $\mathcal{F}T$ by

$$\mathcal{F}T[\chi] = 2\pi\, T[\mathcal{F}^{-1}\chi] \tag{8.36}$$

so as to agree with (8.35) in the restricted case. On the *lhs* here we have the Fourier transformed distribution $\mathcal{F}T$ acting on some test function $\chi \in \mathcal{S}(\mathbb{R})$, whilst on the *rhs* we have the original distribution T acting on the inverse Fourier transform of χ. Again, this definition will be compatible with the idea of obtaining tempered distributions as the limit of a sequence of test functions.

Let's now see how this definition works in practice.

8.3.1 Fourier transform of the Dirac δ

We've seen that, if we want to solve a driven pde $\mathcal{L}y(x) = f(x)$ then its often useful to first construct a Green's function $G(x, \xi)$ obeying $\mathcal{L}G(x, \xi) = \delta(x - \xi)$. Thus, if we wish to use Fourier transforms to solve such equations, we'll need to understand the Fourier transform of distributions such as the δ-function.

In this case, it's easy to obtain the transform naively using the heuristic object $\delta(x)$. We simply compute

$$\mathcal{F}[\delta(x)] = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\delta(x)\,\mathrm{d}x = 1\,. \tag{8.37}$$

Let's repeat this using our general definition for Fourier transforms of distributions. According to (8.36), for any test (Schwartz) function ϕ we have

$$\mathcal{F}\delta[\phi] = 2\pi\,\delta[\mathcal{F}^{-1}\phi] \tag{8.38}$$

and using the definition of the Dirac δ acting on $\mathcal{F}^{-1}\phi$ (thought of as a function of x), we obtain

$$2\pi\,\delta[\mathcal{F}^{-1}\phi] = 2\pi\,\delta\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\,\phi(k)\,\mathrm{d}k\right] = \int_{-\infty}^{\infty} \phi(k)\,\mathrm{d}k = (1,\phi) \tag{8.39}$$

extracting the value of $\mathcal{F}^{-1}\phi$ at $x=0$. The net result is just the inner product of the test function ϕ with the constant function 1. Recalling the definition of distributions that also happen to be ordinary functions, we see that the Fourier transform of the Dirac δ is just 1!

The Fourier inversion theorem assures us that the Fourier transform of 1 itself should be 2π times a δ-function. However, this claim is not obvious, since it's far from clear what we should make of the integral $\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x$. To make progress, we treat 1 as a tempered distribution T_1. Then according to (8.36), the transformed distribution $\mathcal{F}T_1$ is determined by

$$\mathcal{F}T_1[\psi] = 2\pi\,T_1[\mathcal{F}^{-1}\psi] = 2\pi\,(1,\mathcal{F}^{-1}\psi) \tag{8.40}$$

where the second equality recognizes that T_1 is the distribution associated to an ordinary function, 1. Thus, to understand the Fourier transform of 1 we only need take the inverse Fourier transform of a Schwartz function ψ, which is always well–defined. Comparing with (8.38) we have indeed

$$\mathcal{F}T_1[\psi] = 2\pi\,(1,\mathcal{F}^{-1}\psi) = 2\pi\,\delta[\mathcal{F}^{-1}\psi]\,. \tag{8.41}$$

Thus, in the world of tempered distributions, we have

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x = 2\pi\,\delta(k) \tag{8.42}$$

and the Fourier transform of 1 is indeed 2π times the Dirac δ.

The translation and rephasing properties of the Fourier transform provide us with the simple corollaries

$$\mathcal{F}[\delta(x-a)] = \mathrm{e}^{-\mathrm{i}ka} \qquad \text{and} \qquad \mathcal{F}[\mathrm{e}^{-\mathrm{i}\ell x}] = 2\pi\,\delta(k-\ell)\,, \tag{8.43}$$

whilst we also have

$$\begin{aligned} \mathcal{F}[\cos(\ell x)] &= \pi\left(\delta(k+\ell)+\delta(k-\ell)\right) \\ \mathcal{F}[\sin(\ell x)] &= \mathrm{i}\pi\left(\delta(k+\ell)-\delta(k-\ell)\right)\,. \end{aligned} \tag{8.44}$$

In particular, a highly localised signal in physical space (such as a δ function) has a very spread out representation in Fourier space. Conversely, a highly spread out (yet periodic) signal in physical space (such as a sine wave) is highly localised in Fourier space. This is further illustrated in problem sheet 3, where you compute the Fourier transform of a Gaussian. Going further, the fact that Fourier transforms turns multiplication by x into differentiation *wrt* k shows that the Fourier transform of a polynomial involves derivatives of δ-functions, and conversely the Fourier transform of derivatives of δ-functions is a polynomial.

For another example, consider the Heaviside step function

$$\Theta(x) = \begin{cases} 1 & x>0 \\ 0 & \text{else.} \end{cases} \tag{8.45}$$

We first note that $\Theta(x) = \lim_{\epsilon\to 0^+} \Theta(x)\,\mathrm{e}^{-\epsilon x}$ and that for any $\epsilon > 0$

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\,\Theta(x)\,\mathrm{e}^{-\epsilon x}\,\mathrm{d}x = \int_{0}^{\infty} \mathrm{e}^{-(\epsilon+\mathrm{i}k)x}\,\mathrm{d}x = \frac{1}{\epsilon+\mathrm{i}k} \tag{8.46}$$

so that, naively taking $\epsilon \to 0$, we might think that the Fourier transform of $\Theta(x)$ is just $1/\mathrm{i}k$. However, the presence of ϵ was clearly important to ensure convergence of the integral as $x \to \infty$, so to really understand what this transform means we must treat $\mathcal{F}\Theta$ as a distribution and see how it acts on a test function ϕ. For finite ϵ and any small $\delta > 0$ we have

$$\begin{aligned}\mathcal{F}\Theta[\phi] &= \lim_{\epsilon\to 0^+}\int_{-\infty}^{\infty}\frac{1}{\epsilon+\mathrm{i}k}\,\phi(k)\,\mathrm{d}k\\ &= \int_{|k|>\delta}\frac{1}{\mathrm{i}k}\,\phi(k)\,\mathrm{d}k + \lim_{\epsilon\to 0^+}\int_{-\delta}^{\delta}\left[\frac{\phi(k)-\phi(0)}{\epsilon+\mathrm{i}k} + \frac{\phi(0)}{\epsilon\ +\mathrm{i}k}\right]\mathrm{d}k\end{aligned} \tag{8.47}$$

where it is clearly safe to take $\epsilon \to 0$ whenever $|k| > \delta > 0$. For the remaining terms, as $\delta \to 0$ we have

$$\int_{-\delta}^{\delta}\frac{\phi(k)-\phi(0)}{\epsilon+\mathrm{i}k}\,\mathrm{d}k = \frac{1}{i}\int_{-\delta}^{\delta}\frac{d\phi}{dk}\,\mathrm{d}k + \ \mathcal{O}(\epsilon) = \mathcal{O}(\epsilon) \tag{8.48}$$

since the test function ϕ is smooth, and finally

$$\begin{aligned}\lim_{\epsilon\to 0^+}\int_{-\delta}^{\delta}\frac{\phi(0)}{\epsilon+\mathrm{i}k}\,\mathrm{d}k &= -\mathrm{i}\phi(0)\lim_{\epsilon\to 0^+}\ln(\epsilon+\mathrm{i}k)\Big|_{-\delta}^{\delta}\\ &= \pi\phi(0)\,.\end{aligned} \tag{8.49}$$

Combining the pieces, altogether we have

$$\mathcal{F}\Theta[\phi] = \pi\phi(0) + \lim_{\delta\to 0}\int_{|k|>\delta}\frac{1}{\mathrm{i}k}\,\phi(k)\,\mathrm{d}k \tag{8.50}$$

showing that, as a distribution, the Fourier transform of $\Theta(x)$ is

$$\mathcal{F}[\Theta(x)] = \begin{cases}\dfrac{1}{\mathrm{i}k} & k \neq 0\\ \pi\delta(k) & k = 0\,.\end{cases} \tag{8.51}$$

This is sometimes written as $\mathcal{F}[\Theta(x)] = \text{p.v.}\,(\mathrm{i}k)^{-1} + \pi\delta(k)$ where the letters p.v. stand for the (Cauchy) *principal value* and mean that we should exclude the point $k = 0$ from any integral containing this term. What happens at $k = 0$ is instead governed by the δ-function.

8.4 Linear systems and transfer functions

Fourier transform are often used in the systematic analysis of linear systems which arise in many applications. Suppose we have a linear operator $\mathcal{L}$ acting on input $I(t)$ to give output $O(t)$. For example, $\mathcal{L}$ may represent the workings of an amplifier that can modify both the amplitude and phase of individual frequency components of an input signal.

We first write the input signal in terms of its Fourier transform $\tilde{I}(\omega)$:

$$I(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega t}\, \tilde{I}(\omega)\, \mathrm{d}\omega$$

via the Fourier inversion theorem. This expresses the input as a linear combination of components with various definite frequencies ω, with the amplitude and phase of each component given respectively by the modulus and argument of $\tilde{I}(\omega)$. In this context, it is sometimes called the *synthesis* of the input, while the Fourier transform $\tilde{I}(\omega) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\omega t}\, I(\omega)\, \mathrm{d}\omega$ itself is known as the *resolution* of the pulse (into its frequency components).

Now suppose that $\mathcal{L}$ modifies the amplitudes and phases via a complex function $\tilde{R}(\omega)$ to produce the output

$$O(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega t}\, \tilde{R}(\omega)\, \tilde{I}(\omega) \mathrm{d}\omega\,. \tag{8.52}$$

$\tilde{R}(\omega)$ is called the *transfer function* of the system and its inverse Fourier transform $R(t)$ is called the *response function*[34]. Thus the transfer and response functions are related via

$$\tilde{R}(\omega) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega t}\, R(t)\, \mathrm{d}t \qquad \text{and} \qquad R(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega t}\, \tilde{R}(\omega)\, \mathrm{d}\omega\,.$$

Equation (8.52) shows that $R(t)$ is the output $O(t)$ of the system when the input has $\tilde{I}(\omega) = 1$ — in other words when the input signal is $I(t) = \delta(t)$. Hence the response function $R(t)$ is closely related to the Green's function when $\mathcal{L}$ is a differential operator.

According to (8.52) and the convolution theorem, the output is

$$O(t) = \mathcal{F}^{-1}[\tilde{R}(\omega)\tilde{I}(\omega)] = R * I(t) = \int_{\mathbb{R}} R(t-u)\, I(u)\, \mathrm{d}u \tag{8.53}$$

Causality (and the absence of a background hum!) implies that if the input signal vanishes for all $t < t_0$ then there can be no output signal before t_0 either. Since the response function $R(t)$ is the output when the input $I(t) = \delta(t)$, and this input certainly vanishes for all $t < 0$, we see that $R(t) = 0$ for all $t < 0$. If we now assume that the was no input signal before $t = 0$ then (8.53) becomes

$$O(t) = \int_{-\infty}^{\infty} R(t-u)\, I(u)\, \mathrm{d}u = \int_{0}^{t} R(t-u)\, I(u)\, \mathrm{d}u \tag{8.54}$$

so that the output signal is formally the same as we found in our previous expressions in section 7.5 with Green's functions for initial value problems.

8.4.1 General form of transfer functions for ode's

We now consider the case with the relation linear finite order o.d.e. $\mathcal{L}_m I(t) = \mathcal{L}_n O(t)$, where $\mathcal{L}_m$ and $\mathcal{L}_n$ are the differential operators

$$\mathcal{L}_m = \sum_{j=0}^{m} b_j \frac{d^j}{dt^j} \qquad \text{and} \qquad \mathcal{L}_n = \sum_{i=0}^{n} a_i \frac{d^i}{dt^i}$$

[34]Warning! In various texts both $R(t)$ and $\tilde{R}(\omega)$ are referred to as the response or transfer function, in different contexts.

with constant coefficients (a_i, b_j). For simplicity, we'll consider the case $m = 0$ so that the *input* acts directly as a forcing term, but the fully general case is not much more difficult. Taking Fourier transforms we obtain

$$\tilde{I}(\omega) = \left[\sum_{i=0}^{n} a_i(\mathrm{i}\omega)^n\right]\tilde{O}(\omega) \tag{8.55}$$

and therefore

$$\tilde{R}(\omega) = \frac{1}{[a_0 + a_1\mathrm{i}\omega + \cdots + a_n(\mathrm{i}\omega)^n]}. \tag{8.56}$$

Thus the transfer function is a rational function with an n^{th} degree polynomial in $(\mathrm{i}\omega)$ as the denominator. By the fundamental theorem of algebra, this polynomial will have n complex roots, counted with multiplicity. Assuming these roots are at point $c_j \in \mathbb{C}$ and allowing for repeated roots, we write

$$\tilde{R}(\omega) = \frac{1}{a_n}\prod_{j=1}^{J}\frac{1}{(\mathrm{i}\omega - c_j)^k_j} \tag{8.57}$$

where k_j is the multiplicity of the root at c_j, and $\sum_{j=1}^{J} k_j = n$. Using partial fractions, this can be written as a sum of terms of the form

$$\frac{\Gamma_{mj}}{(\mathrm{i}\omega - c_j)^m} \qquad \text{where } 1 \leq m \leq k_j \tag{8.58}$$

where Γ_{mj} are constants so that

$$\tilde{R}(\omega) = \sum_{j=1}^{J}\sum_{m=1}^{k_j}\frac{\Gamma_{mj}}{(\mathrm{i}\omega - c_j)^m}\,. \tag{8.59}$$

To find the response function $R(t)$, we need to compute the inverse Fourier transform of this function.

Let's define $\tilde{h}_m(\omega) := 1/(\mathrm{i}\omega - \alpha)^{m+1}$ with $m \geq 0$; by linearity of the Fourier transform, our job is done once we compute $\mathcal{F}^{-1}[\tilde{h}_m]$. We'll cheat. I'll give you the answer and we'll check it's correct by computing its Fourier transform[35]. Consider the function

$$h_0(t) := \begin{cases} \mathrm{e}^{\alpha t} & t > 0 \\ 0 & \text{else.} \end{cases} \tag{8.60}$$

Provided $\mathrm{Re}(\alpha) < 0$ the Fourier transform is easily computed to be

$$\tilde{h}_0(\omega) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega t}\,h_0(t)\,\mathrm{d}t = \int_0^{\infty} \mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{e}^{\alpha t}\,\mathrm{d}t = \frac{1}{\mathrm{i}\omega - \alpha} \tag{8.61}$$

[Note that if $\mathrm{Re}(\alpha) > 0$ then the above integral is certainly divergent. Indeed recalling the theory of linear constant coefficient ode's, we see that the c_j's above are the roots of the

[35]Going in the 'forwards' direction is most simply done with some complex analysis you'll meet next term, but you can also do it by differentiating under the integral if you're careful.

auxiliary equation and the equation has solutions with terms $e^{c_j t}$ which grow exponentially if $\mathrm{Re}(c_j) > 0$. Such exponentially growing functions are problematic for Fourier theory, so here we will consider only the case $\mathrm{Re}(c_j) < 0$ for all j, corresponding to 'stable' odes whose solutions do not grow unboundedly as $t \to \infty$. The whole subjects of stability and control grow from these remarks.]

Next consider the function

$$h_1(t) := \begin{cases} te^{\alpha t} & t > 0 \\ 0 & \text{else,} \end{cases} \tag{8.62}$$

so $h_1(t) = th_0(1)$. Recalling that 'multiplication becomes differentiation' for Fourier transforms (so $\mathcal{F}[xf(x)] = \mathrm{i}d\mathcal{F}[f]/dk$), we obtain

$$\tilde{h}_1(\omega) = \mathrm{i}\frac{d}{d\omega}\frac{1}{\mathrm{i}\omega - \alpha} = \frac{1}{(\mathrm{i}\omega - \alpha)^2} \tag{8.63}$$

(which may also be derived directly by evaluating the $\int_0^\infty te^{\alpha t}e^{-iwt}\,dt$). Similarly (or using proof by induction), for $\mathrm{Re}(\alpha) < 0$ we have

$$h_m(t) = \begin{cases} \dfrac{t^m}{m!}e^{\alpha t} & t > 0 \\ 0 & \text{else} \end{cases} \qquad \text{has Fourier transform} \qquad \tilde{h}_m(\omega) = \frac{1}{(\mathrm{i}\omega - \alpha)^{m+1}} \tag{8.64}$$

with $m \geq 0$. Thus, for such stable systems, we can use this in (8.54) to easily construct the output from the input. Notice that all the $h_m(t)$ decay as $t \to \infty$, but after $t = 0$ they can initially increase to some finite time maximum (at $t = t_m = m/|\alpha|$ if $\alpha < 0$ and real for example). We also see that the response function $R(u) = 0$ for $u < 0$ so that, as expected, the upper limit in (8.53) should be t and the output $O(t)$ depends only on the input at earlier times.

To see more clearly the relation of above formalism to Green's functions, let's consider the familiar equation

$$\frac{d^2y}{dt^2} + 2p\frac{dy}{dt} + (p^2 + q^2)y = f(t) \tag{8.65}$$

which, provided $p > 0$, describes the motion of a forced, damped oscillator. Since $p > 0$, the drag force $-2py'$ acts opposite to the direction of velocity so the motion is damped. We assume that the forcing term $f(t)$ is zero for $t < 0$. Also $y(t)$ and $y'(t)$ are also zero for $t < 0$ and we have initial conditions $y(0) = y'(0) = 0$. The Fourier transformed equation is $(\mathrm{i}\omega)^2\tilde{y} + 2\mathrm{i}p\omega\tilde{y} + (p^2 + q^2)\tilde{y} = \tilde{f}$ and so

$$\tilde{y} = \frac{\tilde{f}}{-\omega^2 + 2\mathrm{i}p\omega + (p^2 + q^2)} =: \tilde{R}(\omega)\,\tilde{f}(\omega) \tag{8.66}$$

which solves the equation algebraically in Fourier space. To find the solution in real space we take the inverse Fourier transform to find

$$y(t) = \int_0^t R(t-u)\,f(u)du = \int_0^t \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{\mathrm{i}\omega(t-u)}}{p^2 + q^2 + 2\mathrm{i}p\omega - \omega^2}\,\mathrm{d}\omega\right] f(u)\,\mathrm{d}u\,. \tag{8.67}$$

where the quantity in square brackets is the response function $R(t-u)$. Now consider $\mathcal{L}R(t-u)$, using this integral formulation. Assuming that formal differentiation within the integral sign is valid we have

$$\begin{aligned}
&\frac{d^2}{dt^2}R(t-u) + 2p\frac{d}{dt}R(t-u) + (p^2+q^2)R(t-u) \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega(t-u)} \left[\frac{(\mathrm{i}\omega)^2 + 2\mathrm{i}p\omega + (p^2+q^2)}{p^2+q^2+2\mathrm{i}p\omega - \omega^2}\right] \mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega(t-u)}\,\mathrm{d}\omega = \delta(t-u),
\end{aligned} \tag{8.68}$$

where we used that 1 is the Fourier transform of the δ-function. Therefore, the Green's function $G(t,u)$ is simply the response function $R(t-u)$ by (mutual) definition. In the problem sheets you're asked to fill in the details of this example, computing both $R(t)$ and the Green's function explicitly.

Fourier transforms can also be used to solve ode's on the full line $\mathbb{R}$, so long as the functions involved have sufficiently good asymptotic properties for the Fourier integrals to exist. As an example consider suppose $y : \mathbb{R} \to \mathbb{R}$ solves the ode

$$\frac{d^2y}{dx^2} - A^2 y = -f(x) \tag{8.69}$$

and is such that $y \to 0$ and $y' \to 0$ as $|x| \to \infty$. We take A to be a positive real constant. Taking the Fourier transform, we can solve the equation in Fourier space trivially to find

$$\tilde{y}(k) = \frac{\tilde{f}(k)}{A^2+k^2} \tag{8.70}$$

and so to recover the solution $y(x)$ we need to take the inverse Fourier transform of the product $\tilde{f}(k) \times 1/(A^2+k^2)$. From the convolution theorem, this will be the convolution of $f(t)$ with a function whose Fourier transform is $1/(A^2+k^2)$. Consider the function

$$h(x) = \frac{\mathrm{e}^{-\mu|x|}}{2\mu} \qquad \text{where} \qquad \mu > 0\,. \tag{8.71}$$

Since $h(x)$ is even, its Fourier transform can be written as

$$\tilde{h}(k) = \mathrm{Re}\left(\frac{1}{\mu}\int_0^{\infty} \exp\left[-(\mu+\mathrm{i}k)x\right]\mathrm{d}x\right) = \frac{1}{\mu}\mathrm{Re}\left(\frac{1}{\mu+\mathrm{i}k}\right) = \frac{1}{\mu^2+k^2}\,, \tag{8.72}$$

which identifies $h(x)$ as the function we seek. The convolution theorem then gives

$$y(x) = \frac{1}{2A}\int_{-\infty}^{\infty} \mathrm{e}^{-A|x-u|}\,f(u)\,\mathrm{d}u\,. \tag{8.73}$$

This solution is clearly in the form of a Green's function expression. Indeed the same expression may be derived using the Green's function formalism of chapter 7, applied to the infinite domain $(-\infty,\infty)$ and imposing suitable asymptotic boundary conditions on the Green's function for $|x| \to \infty$.

8.5 The discrete Fourier Transform

In realistic, physical situations or in numerical applications, we cannot add expect to know the value of $f(x)$ for *every* $x \in \mathbb{R}$, but only within some finite interval $x \in [-R, S]$. However, provided $|f(x)|$ is small outside this range, we may reasonably expect to approximate the Fourier transform by

$$\tilde{f}(k) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} f(x)\,\mathrm{d}x \approx \int_{-R}^{S} \mathrm{e}^{-\mathrm{i}kx} f(x)\,\mathrm{d}x\,. \tag{8.74}$$

Even within this range we cannot numerically evaluate this integral by computing $\mathrm{e}^{-\mathrm{i}kx} f(x)$ at *every* value of x, but will rather have to sample it at some finite set of points. For simplicity, let's assume these points are equally spaced, at $x_j = -R + j(R+S)/N$ for $j = 0, \ldots, N-1$ where N is some large positive integer. Then we can set

$$\tilde{f}(k) \approx \frac{R+S}{N} \sum_{j=0}^{N-1} \mathrm{e}^{-\mathrm{i}kx_j} f(x_j)\,. \tag{8.75}$$

Similarly, although the Fourier transform was originally valid for all $k \in \mathbb{R}$, in practice our computer's memory can only store the values of $\tilde{f}(k)$ for some finite set values of k. It's convenient to choose these to be at $k = k_m := 2\pi m/(R+S)$ for some integer(s) m. Then we have

$$\tilde{f}(k_m) \approx \frac{(R+S)\,\mathrm{e}^{\mathrm{i}k_m R}}{N} \sum_{j=0}^{N-1} \mathrm{e}^{-2\pi\mathrm{i}jm/N} f(x_j) = \frac{(R+S)\,\mathrm{e}^{\mathrm{i}k_m R}}{N} \sum_{j=0}^{N-1} f(x_j)\,\omega^{-mj} \tag{8.76}$$

where $\omega = \mathrm{e}^{2\pi\mathrm{i}/N}$ in an N^{th} root of unity.

Thus, for numerical purposes, computing Fourier transforms amounts to computing expressions of the form $F(m) := \frac{1}{N}\sum_{j=0}^{N-1} f(x_j)\,\omega^{-mj}$. Of course, unlike the exact Fourier transform, these truncations have thrown away much of the information about our function: we obviously cannot see structures that vary on scales shorter than our sampling interval $(R+S)/N$. Furthermore, since it is built from N^{th} roots of unity, evidently our approximate function obeys $F(m+N) = F(m)$ whereas our original function $f(x)$ and its exact Fourier transform $\tilde{f}(k)$ were not assumed to obey any periodicity conditions[36]. Nevertheless, we shall see that knowing the values of the *sums* $F(m)$ for all $m = 0, 1, \ldots, N-1$ is sufficient to reproduce the exact values of $f(x)$ at all the points x_k.

To proceed, let G be the set $\{1, \omega, \omega^2, \ldots, \omega^{N-1}\}$ of N^{th} roots of unity, and let $f : G \to \mathbb{C}$ and $g : G \to \mathbb{C}$ be functions on G. The space of such functions is a complex vector space, and we can define an inner product on this space by

$$(f, g) = \frac{1}{N} \sum_{j=0}^{N-1} f(\omega^j)^*\, g(\omega^j) \tag{8.77}$$

[36]In fact, if $|f(x)| \to 0$ sufficiently rapidly as $|x| \to \infty$ that we are justified in ignoring $x \notin [-R, S]$ then $\tilde{f}(k)$ will also decay as $m \to \infty$.

in analogy to our earlier treatment of functions on $\mathbb{R}$. It is a straightforward exercise to check that this inner product obeys the usual sesquilinearity properties

$$(f, c_1 g_1 + c_2 g_2) = c_1(f, g_2) + c_2(f, g_2) \qquad \text{and} \qquad (g, f)^* = (f, g)$$

for constants $c_{1,2}$, and that $(f, f) \geq 0$ with equality iff $f(\omega) = 0$ identically. Next, we claim that the functions $e_m(\omega^j) := \omega^{mj}$ obey $(e_m, e_n) = \delta_{mn}$ and so form a set of basis functions on G that are orthonormal wrt (,). To see this, note that

$$(e_m, e_m) = \frac{1}{N} \sum_{j=0}^{N-1} \omega^{-mj} \, \omega^{mj} = \frac{1}{N} \sum_{j=0}^{N-1} 1 = 1 \tag{8.78}$$

whilst whenever $m \neq n$ we have

$$(e_n, e_m) = \frac{1}{N} \sum_{j=0}^{N-1} \omega^{(m-n)j} = \frac{1}{N} \frac{\omega^{(m-n)N} - 1}{\omega^{(m-n)} - 1} = 0 \tag{8.79}$$

since $\omega^{(m-n)N} = 1$ for $m \neq n$ and ω an N^{th} root of unity.

These results provide us with a Fourier expansion for functions on G. We can expand

$$f(\omega^m) = \sum_{j=0}^{N-1} \hat{f}_j \, e_m(\omega^j) \qquad \text{where} \qquad \hat{f}_k \equiv (e_k, f) = \frac{1}{N} \sum_{k=0}^{N-1} \omega^{-jk} \, f(\omega^k) \, . \tag{8.80}$$

In fact, since G here is just a finite set, there are no issues about convergence to worry about, and the validity of this expansion is easily proved using the orthonormality of the $\{e_m\}$. Likewise, Parseval's theorem is easily proved by writing

$$\begin{aligned}(f, f) &= \frac{1}{N} \sum_{j=0}^{N-1} f(\omega^j)^* \, f(\omega^j) = \frac{1}{N} \sum_{j=0}^{N-1} \left(\sum_{l,m=0}^{N-1} \hat{f}_m^* \, \hat{f}_l \; e_m(\omega^j)^* \, e_l(\omega^j) \right) \\ &= \sum_{l,m} \hat{f}_m^* \, \hat{f}_l \left(\sum_{j=0}^{N-1} e_m(\omega^j)^* \, e_l(\omega^j) \right) = \sum_{l,m} \hat{f}_m^* \, \hat{f}_l \, \delta_{ml} = \sum_{m=0}^{N-1} \hat{f}_m^* \, \hat{f}_m = N(\hat{f}, \hat{f}) \, ,\end{aligned} \tag{8.81}$$

with no complications about exchanging the order of the finite sums. Again the Fourier transform preserves the inner product.

Now we can understand why knowledge of the sums

$$F(m) = \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) \, \omega^{-mj}$$

for all m is sufficient to reproduce the exact values of the sampled function $f(x)$ at the sampling points $x = \{x_0, x_1, \ldots, x_{N-1}\}$. Defining a function $g : G \to \mathbb{C}$ by $g(\omega^j) := f(x_j)$, we have $F(m) = (e_m, g)$ and therefore

$$f(x_j) = \sum_{m=0}^{N-1} \omega^{jm} \, F(m) \tag{8.82}$$

so that the $f(x_j)$ can be recovered exactly from the $F(m)$.

Since the discrete Fourier transform acts linearly on the space of functions from $G \to \mathbb{C}$, and G is a finite set, we can represent it by a matrix. Let $\mathbf{U}$ be the $N \times N$ symmetric matrix with entries $U_{mj} = e_m(\omega^j) = \omega^{mj} = e_j(\omega^m)$ with $m, j = 0, 1, 2, \ldots, N-1$. Then if we let $\mathbf{f}$ be the (column) vector $\mathbf{f} = (f(x_0), f(x_1), \ldots, f(x_{N-1}))^{\mathrm{T}}$ the discrete Fourier transform may be written as the matirx multiplication

$$\hat{\mathbf{f}} = \mathbf{U}\mathbf{f}\,. \tag{8.83}$$

The matrix $\mathbf{U}/\sqrt{N}$ is in fact *unitary*, so $\mathbf{U}^{-1} = \mathbf{U}^\dagger/N$. To see this, note that $\mathbf{U}^\dagger$ has entries $U^\dagger_{jk} = e_k(\omega^j)^* = \omega^{-jk}$ and so

$$(\mathbf{U}\mathbf{U}^\dagger)_{mk} = \sum_{j=0}^{N-1} e_m(\omega^j)\, e_k^*(\omega^j) = N(e_k, e_m) = N\delta_{mk} \tag{8.84}$$

from the orthonormality properties of the basis functions $\{e_m\}$. Thus we also have

$$\mathbf{f} = \mathbf{U}^{-1}\hat{\mathbf{f}} = \frac{1}{N}\mathbf{U}^\dagger\hat{\mathbf{f}} \tag{8.85}$$

for the inverse transform.

Hopefully, you've been awake enough throughout this course to notice that the Fourier series for functions[37] $f : T^n \to \mathbb{C}$, Fourier transforms for functions $f : \mathbb{R}^n \to \mathbb{C}$ and discrete Fourier transform for functions $f : G \to \mathbb{C}$ have all looked rather similar. In particular, in each case we have an Abelian group $\mathcal{G}$ (either T^n, $\mathbb{R}^n$ or G) and maps $e_{\mathrm{k}} : \mathcal{G} \to \mathbb{C}$, given on T^n and $\mathbb{R}^n$ by $\exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x})$, with $\mathbf{k} \in \mathbb{Z}^n$ and $\mathbf{k} \in \mathbb{R}^n$ respectively, and by $e_k(\omega^j) = \omega^{jk}$ on G. These functions all obey

$$e_{\mathbf{k}}(\mathbf{x} + \mathbf{y}) = e_{\mathbf{k}}(\mathbf{x})\, e_{\mathbf{k}}(\mathbf{y}) \tag{8.86}$$

and $e_{\mathbf{k}}(I) = 1$ where I is the identity element of the group $\mathcal{G}$ ($\mathbf{0}$ in $\mathbb{R}^n$, $\mathbf{1}$ in T^n and 1 in G). In fact, it's not much more work to establish Fourier theory for any Abelian group, though we won't do it here. There's also a powerful generalization of Fourier theory to non–Abelian groups, at least for compact groups, which you can read about here.

8.5.1 Fast Fourier transforms

Non-examinable. All the best bits are.

In the months that followed the Cuban missile crisis, geopolitical tensions were running high. The realization of just how close the world had come to a nuclear apocalypse persuaded many that a curb on the development of nuclear weapons was required, and negotiations were underway to work towards a Nuclear Test Ban Treaty. A major problem was that neither side trusted the other to stick to the terms of the Treaty. In particular,

[37]Here $T^n = S^1 \times S^1 \times \cdots \times S^1$ is an n-dimensional torus. Equivalently, we have considered this case when Fourier analysing functions periodic in n variables.

the US wanted to be sure they could tell whether or not a nuclear test had taken place deep within the Soviet Union without having to undertake politically sensitive visits to the actual site. Overground tests were relatively easy to spot, not least because of the fallout effects, but detecting underground tests was another matter.

During a meeting of President Kennedy's Scientific Advisory Committee, John Tukey of Princeton got interested in a proposal to place many sensitive seismological devices at locations surrounding the Soviet Union. These devices could monitor vibrations, recording this information over time. But in order to tell whether a test had occurred, one would have to separate out the signal of the explosion from other noises and rumblings of the Earth's crust. This called for Fourier analysis, which would separate out the frequencies of vibration, but the large landmass of Soviet Russia and the long period of time for which monitoring would be required meant that a powerful computer would be needed to perform the analysis. The problem is that to actually compute $\hat{f}(m)$ at even a single value of m using the formula

$$\hat{f}(m) \approx \frac{1}{N} \sum_{j=0}^{N-1} \omega^{-mj} f(x_j) \tag{8.87}$$

obtained above is very computationally expensive. Even if our computer is provided with a library containing the values of ω^{-mj}, the $f(x_j)$ will depend on what the seismometers record, so we shall need at least $N-1$ additions and $N+1$ multiplications: a total of $2N$ operations to compute $\hat{f}(m)$ for a single value of m. Thus, even given all the values of ω^{-mj}, computing $\hat{f}(m)$ for all $m = 0, \ldots, N-1$ this way will thus take $2N^2$ operations.

Together with James Cooley at IBM, Tukey realised that a much more efficient algorithm could be obtained by breaking the task down into smaller chunks. Suppose that $N = 2M$ for some integer M. Then for a function on the $N = 2M^{\text{th}}$ roots of unity we can write

$$\begin{aligned}\frac{1}{N} \sum_{j=0}^{N-1} \omega^{-mj} g(\omega^j) &= \frac{1}{2M} \sum_{j=0}^{2M-1} \omega^{-mj} g(\omega^j) \\ &= \frac{1}{2}\left[\frac{1}{M} \sum_{k=0}^{M-1} \omega^{-2kj} g(\omega^{2k}) + \frac{1}{M} \sum_{k=0}^{M-1} \omega^{-(2k+1)m} g(\omega^{2k+1})\right]\end{aligned} \tag{8.88}$$

where in going to the second line we've separated out the even ($j = 2k$) and odd ($j = 2k+1$) terms in the sum. Now let η denote an M^{th} root of unity and define $G(\eta^k) \equiv g(\omega^{2k})$ and $H(\eta^k) \equiv g(\omega^{2k+1})$. Then we can write the *rhs* of (8.88) as

$$\frac{1}{2}\left[\frac{1}{M} \sum_{k=0}^{M-1} \eta^{-mk} G(\eta^k) + \frac{\omega^{-m}}{M} \sum_{k=0}^{M-1} \eta^{-mk} H(\eta^k)\right] = \frac{1}{2}\left[\hat{G}(m) + \omega^{-m}\hat{H}(m)\right], \tag{8.89}$$

where the final equality identifies the sums as being the discrete Fourier transforms of G and H as functions on the M^{th} roots of unity.

Given the values of $\hat{G}(m)$, $\hat{H}(m)$ and ω^{-m} for some specific m, these expressions show that we can compute the value of $\hat{f}(m)$ using no more than one addition and two multiplications. Consequently, given all these values, we can compute $\hat{f}(m)$ for *all* $m \in \{0, 1, \ldots, N-1\}$ using at most $3 \times 2M = 6M$ operations. Suppose it takes no more that P_M operations to compute the discrete Fourier transform of any function on the M^{th} roots of unity. Computing the value of ω^{-m} for every $0 \leq m < 2M$ will cost no more than $2M$ multiplications, and so we will be able to compute the value of $\hat{f}(m)$ at *every* $m \in \{0, 1, \ldots, N-1\}$ using no more than $P_{2M} = 8M + 2P_M$ operations.

Now to the point. Let's prove inductively that when $N = 2^n$, $P_N \leq 2^{n+2}n = 4N \log_2 N$ — note that for $N \gg 1$ this is a vast saving over the $\sim 2N^2$ operations required by a 'direct' approach. When $n = 1$ the Fourier transform is completely given by

$$
\begin{aligned}
\hat{F}(1) &= \frac{1}{2}\left(F(1) + F(-1)\right) \\
\hat{F}(-1) &= \frac{1}{2}\left(F(1) + (-1)F(-1)\right) ,
\end{aligned}
\tag{8.90}
$$

and these values clearly take no more than 5 operations on $\{F(1), F(-1)\}$ to obtain. Thus, certainly $P_2 = 5 \leq 8$. Assuming the result when $N = 2^m$ for some m, our recurrence then shows that $P_{2^{m+1}} \leq 8 \cdot 2^m + 2 \times 2^{m+2}m = 2^{m+3}(m+1)$ and the induction holds.

Overnight, Cooley and Tukey's algorithm revolutionized our ability to actually compute Fourier transforms numerically. Aside from Nuclear Test Ban Treaties, nowadays their results (and further refinements) lie behind essentially all the image processing computations that enable you see what your friends are up to via iPhone. But they weren't the first to get there. Back in 1805 Gauss had been trying to reconstruct the orbits of comets Pallas and Juno from their various intermittent observations. In the margins of his notebooks, he'd discovered Cooley & Tukey's algorithm for himself as a means to speed up his calculations.