.

1B METHODS LECTURE NOTES

Richard Jozsa, DAMTP Cambridge
rj310@cam.ac.uk
October 2010

PART III:

Inhomogeneous ODEs; Fourier transforms

6 THE DIRAC DELTA FUNCTION

The Dirac delta function and an associated construction of a so-called Green's function will provide a powerful technique for solving inhomogeneous (forced) ODE and PDE problems.

To motivate the introduction of the delta function consider the set $\mathcal{F}$ of (suitably well behaved) real functions $f : \mathbb{R} \to \mathbb{R}$. Any such function g may be viewed as a map G from $\mathcal{F}$ to $\mathbb{R}$ viz. for any f we set

$$G(f) = \int_{-\infty}^{\infty} g(x) f(x)\, dx \tag{1}$$

i.e. g appears as the kernel of an integral with variable f (and here we are just assuming that the integrals exist). Such "functions of a function" are called *functionals*. These functionals are clearly linear in their argument f and we ask: do we get all linear maps from $\mathcal{F}$ to $\mathbb{R}$ in this way? Well, consider the following (clearly linear) functional

$$L(f) = f(0).$$

What should the kernel be? If we call it $\delta(x)$ we would need

$$\int_{-\infty}^{\infty} \delta(x) f(x)\, dx = f(0) \quad \text{for all } f. \tag{2}$$

If we're thinking of δ as being a bonafide continuous function then if $c \neq 0$, we should have $\delta(c) = 0$ – because if $\delta(c) \neq 0$ we could choose an f that is nonzero only very near to c and eq. (2) could be made to fail. Next choosing $f(x) = 1$ we need $\int_{-\infty}^{\infty} \delta(x)\, dx = 1$ yet $\delta(x) = 0$ for all $x \neq 0$! Thus intuitively the graph of $\delta(x)$ must enclose unit area but over a vanishingly small base i.e. we need an "infinite spike" at $x = 0$. Such objects that are not functions in the usual sense but have well defined properties such as eq. (2) are called **generalised functions** or **distributions**. It is possible to develop a mathematically rigorous theory of them and below we'll outline how to do this. But for this "methods" course we will use them only via a set of rules (cf below) motivated intuitively (but the rules can be justified in the rigorous theory).

One possible approach to making sense of the delta function is to view it as the *limit of a sequence* of (ordinary, integrable) functions $P_n(x)$ that should have the property

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} P_n(x) f(x)\, dx = f(0) \tag{3}$$

and we also normalise them to have $\int_{-\infty}^{\infty} P_n(x)\, dx = 1$ for all n. Such sequences are not unique – it can be shown (not obvious in all cases – we'll see eqs. (5,6) in Fourier transforms later!) that each of the following sequences has the required properties:

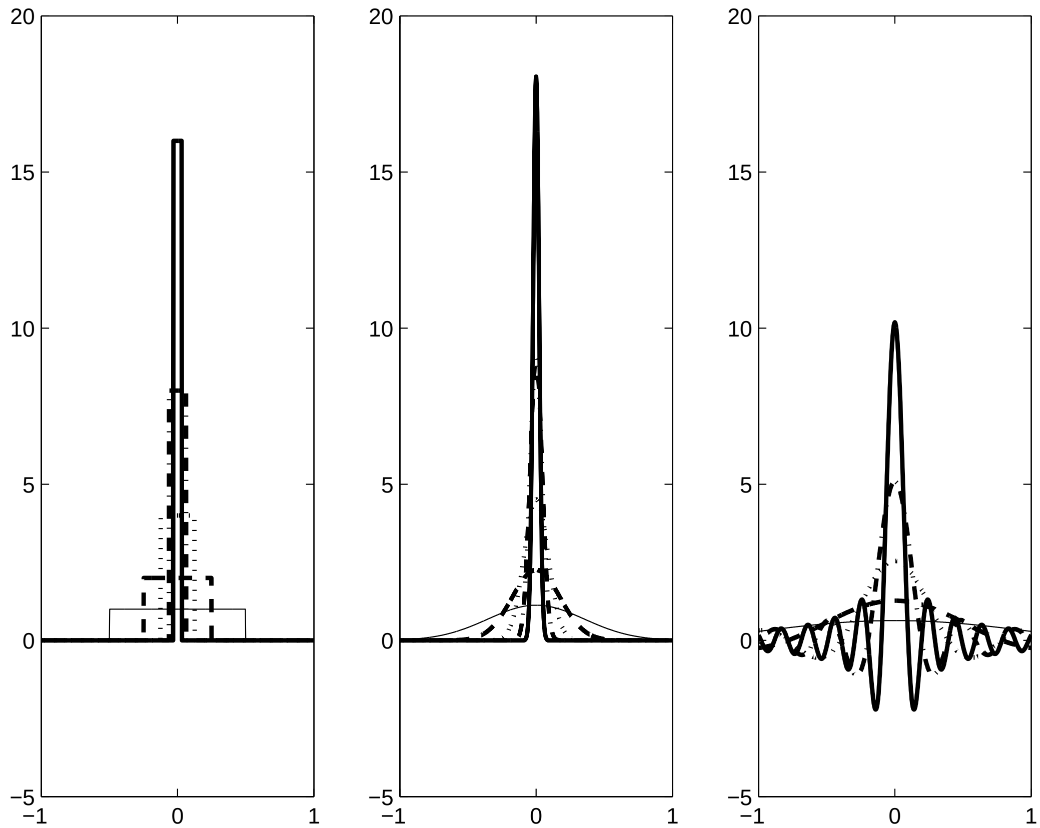

Figure 1: $P_n(x)$ for the top-hat, the Gaussian, and $\sin(nx)/(\pi x)$, showing convergence towards the delta function, for $n = 2, 4, 8, 16, 32$ (thin solid, dashed, dotted, dot-dashed, thick solid).

$$P_n(x) = \begin{cases} \frac{n}{2} & |x| < \frac{1}{n}; \\ 0 & \text{otherwise}, \end{cases} \tag{4}$$

$$P_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2x^2} \tag{5}$$

$$P_n(x) = \frac{\sin(nx)}{\pi x}. \tag{6}$$

These are shown in the figure.

These examples also show that the limit in eq. (3) cannot be taken inside the integral to form a "limit kernel" $\delta(x) \stackrel{?}{=} \lim_{n\to\infty} P_n(x)$ since $P_n(0) \to \infty$ in all cases! Note that the sequence in eq. (6) does not even have $P_n(c) \to 0$ for $c \neq 0$ e.g. $P_n(\pi/2) = \frac{\sin(n\pi/2)}{\pi^2/2}$ which is $\pm 2/\pi^2$ for all odd n. Indeed in eq. (3) contributions from $f(c)$ for $c \neq 0$ are eliminated by P_n oscillating faster and faster near $x = c$ and thus contributing zero effect by increasingly exact cancellations of the oscillations with envelope f for any suitably continuous f.

Physical significance of delta functions

In physics the delta function models *point sources* in a continuum e.g. suppose we have a unit point charge at $x = 0$ (in one dimension). Then its charge density $\rho(x)$ should satisfy $\rho(x) = 0$ for $x \neq 0$ and $\int_{-\infty}^{\infty} \rho(x)\,dx = \text{total charge} = 1$ i.e. $\rho(x) = \delta(x)$ and the physical intuition is well modelled by the sequence eq. (4). In mechanics delta functions model *impulses* e.g. for a particle in one dimension with momentum $p = mv$, Newton's law gives $dp/dt = F$ so $p(t_2) - p(t_1) = \int_{t_1}^{t_2} F\,dt$. If a particle is struck impulsively (e.g. hammer blow) at $t = 0$ within interval (t_1, t_2), with force acting over a vanishingly small time Δt, and resulting in a finite momentum change $\Delta p = 1$ say, then $\int_{t_1}^{t_2} F\,dt = 1$ and F is nonzero only very near $t = 0$. So in the limit of vanishing time interval Δt, $F(t) = \delta(t)$ models the impulsive force. The delta function was introduced by P. A. M. Dirac in the 1930s from considerations of position and momentum in quantum mechanics.

6.1 Properties of the delta function

• Note that in the basic defining property eq. (2) we can take the range of integration to be any interval $[a, b]$ that contains $x = 0$, because we can replace f by the function $\tilde{f}$ that agrees with f on $[a, b]$ and is zero outside $[a, b]$ (or alternatively we can use a sequence such as eq. (4)). If $[a, b]$ does not contain $x = 0$ then the integral is zero.

Properties of the delta function can be intuitively obtained by manipulating the integral in eq. (2) in standard ways, manipulating $\delta(x)$ *as though* it were a genuine function. (You are asked to do some of these on exercise sheet 3). The validity of this procedure can be justified in a rigorous theory of generalised functions.

• Substituting $x' = x - c$ in the integral below we get the so-called **sampling property**:

$$\int_a^b f(x)\delta(x-c)\,dx = \begin{cases} f(c) & a < c < b; \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

(i.e. $\delta(x - c)$ has the spike at $x = c$).

• Substituting $x' = ax$ in the integral with kernel $\delta(ax)$ we get the **scaling property**:

$$\text{if } a \neq 0 \text{ then} \quad \delta(ax) = \frac{1}{|a|}\delta(x)$$

By this equality (and those below) we mean that both sides behave in the same way if used as the kernel of an integral viz. $\int_{-\infty}^{\infty} \delta(ax) f(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{|a|}\delta(x) f(x)\,dx$.

• If $f(x)$ has simple zeroes at n isolated points $x_1, \ldots, x_n$ then $\delta(f(x))$ will have spikes at these x values and

$$\delta(f(x)) = \sum_{k=1}^{n} \frac{\delta(x - x_k)}{|f'(x_k)|}.$$

• If $g(x)$ is integrable and continuous around $x = 0$ then

$$g(x)\delta(x) = g(0)\delta(x).$$

• Using the first bullet point above, we can see that the **integral of the delta function** is the Heaviside step function $H(x)$:

$$\int_{-\infty}^{x} \delta(\xi)\, d\xi = H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x > 0 \end{cases}$$

(and usually we set $H(0) = 1/2$).

More generally, delta functions characteristically appear in derivatives of functions with jump discontinuities e.g. if f on $[a, b]$ has a jump discontinuity of size K at $a < c < b$ and is differentiable for $x < c$ and $x > c$, then $f'(x) = g(x) + K\delta(x - c)$ where $g(x)$ for $x \neq c$ is the derivative of $f(x)$ (and $g(c)$ may be given any value). Indeed integrating this f' to a variable upper limit x, we regain $f(x)$ with the delta function introducing the jump as x crosses $x = c$.

• We can develop the notion of **derivative of the delta function** by formally applying integration by parts to δ' and then using the defining properties of δ itself:

$$\begin{aligned} \int_{-\infty}^{\infty} \delta'(x - c) f(x)\, dx &= [\delta(x - c) f(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \delta(x - c) f'(x)\, dx \\ &= 0 - f'(c) \end{aligned}$$

Similarly, provided $f(x)$ is sufficiently differentiable, we can formally derive:

$$\int_{-\infty}^{\infty} f(x)\delta^{(n)}(x)\, dx = (-1)^n f^{(n)}(0).$$

Example. Compute $I = \int_0^\infty \delta'(x^2 - 1)x^2\, dx$.

The substitution $u = x^2 - 1$ gives $dx = \frac{du}{2\sqrt{u+1}}$ and

$$I = \int_{-1}^{\infty} \delta'(u) \frac{\sqrt{u+1}}{2}\, du = \left[-\frac{d}{du}\left(\frac{\sqrt{u+1}}{2}\right) \right]_{u=0} = -\frac{1}{4}.$$

Towards a rigorous mathematical theory of generalised functions and their manipulation (optional section).

A (complex valued) function ϕ on $\mathbb{R}$ is called a *Schwarz function* if $\phi, \phi', \phi'', \ldots$ are all defined and

$$\lim_{x \to \pm\infty} x^m \phi^{(n)}(x) = 0 \quad \text{for all } m, n = 0, 1, 2, \ldots. \tag{8}$$

(e.g. $\phi(x) = p(x)e^{-x^2}$ for any polynomial $p(x)$). These conditions requiring ϕ to be extremely well behaved asymptotically, will be needed to guarantee the convergence (existence) of integrals from $-\infty$ to ∞ involving modified versions of the ϕ's. Let $\mathcal{S}$ be the set of all Schwarz functions.

Now let g be any continuous function on $\mathbb{R}$ that is "slowly growing" in the sense that

$$\lim_{x \to \pm\infty} \frac{g(x)}{x^n} = 0 \quad \text{for } \textit{some } n = 0, 1, 2, \ldots. \tag{9}$$

(e.g. $x^3 + x$, e^{-x^2}, $\sin x$, $x \ln |x|$ but not e^x, , e^{-x} etc.) Then g defines a functional i.e. linear map, from $\mathcal{S}$ to $\mathbb{C}$ via

$$g\{\phi\} = \int_{-\infty}^{\infty} g(x)\phi(x)\,dx \qquad \text{any } \phi \in \mathcal{S} \tag{10}$$

(noting that the integral is guaranteed to converge by our restrictions on g and ϕ). Next note that the functional corresponding to g' can be related to that of g using integration by parts:

$$g'\{\phi\} = \int_{-\infty}^{\infty} g'(x)\phi(x)\,dx = [g\phi]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} g(x)\phi'(x)\,dx = g\{-\phi'\}.$$

The boundary terms are guaranteed to be zero by eqs. (8,9). Thus

$$g'\{\phi\} = g\{-\phi'\} \tag{11}$$

which is derived (correctly) by standard manipulation of integrals of honest functions. Note that now the action of g' on any ϕ is defined wholly in terms of the action of the *original* g on a *suitable modified* argument viz. $-\phi'$.

Next recall that there are more functionals on $\mathcal{S}$ than those arising from g's as above. For example we argued intuitively that the functional $\delta\{\phi\} \stackrel{\text{def}}{=} \phi(0)$ does not arise from any such g. Now suppose we have *any* given functional $G\{\phi\}$ and we want to establish a notion of derivative G' (or any other modified version of G) for it. The key new idea here is the following procedure: we first consider G's that arise from honest functions g and then express the action of g' (or any other modified version of g) on ϕ, entirely in terms of the action of g itself on a suitable modified argument (as we did above in eq. (11) for the case of derivatives). We then use the latter formula with g replaced by any G as the **definition** of G' i.e. we *define* $G'\{\phi\}$ to be $G\{-\phi'\}$ for *any* G. For example we get

$$\delta'\{\phi\} \stackrel{\text{def}}{=} \delta\{-\phi'\} = -\phi'(0)$$

where the last equality used the (given) definition of δ itself.

The final result is the same as what we would get (non-rigorously) if we "pretended" that $\delta\{\phi\} = \int_{-\infty}^{\infty} \delta(x)\phi(x)\,dx$ for some imagined ("generalised") function $\delta(x)$ on $\mathbb{R}$, and then formally manipulated it the integral $\int_{-\infty}^{\infty} \delta'(x)\phi(x)\,dx$.

Similarly the translation ($x \to x - a$) and scaling ($x \to ax$) and other rules that we listed previously can all be rigorously justified using the above Schwarz function formalism. This formalism also extends in a very natural way to make rigorous sense of Fourier transforms of generalised functions that we'll encounter in chapter 8. If you're interested in learning more about generalised functions, I'd recommend looking at chapter 7 of D. Kammler, "A first course in Fourier analysis" (CUP).

Example - warning. If we think of functionals G as generalised functions γ via $G\{\phi\} = \int_{-\infty}^{\infty} \gamma(x)\phi(x)\,dx \equiv G_\gamma\{\phi\}$ then there is a temptation to add and multiply the γ's just like we do routinely for ordinary functions. For linear combinations this works fine i.e. $G_{c_1\gamma_1 + c_2\gamma_2} = c_1 G_{\gamma_1} + c_2 G_{\gamma_2}$ but *products* are problematic: $G_{\gamma_1}\{\phi\} G_{\gamma_2}\{\phi\} \neq G_{\gamma_1\gamma_2}\{\phi\}$!

This is because the γ's are being used as kernels in integrals and using $\gamma_1\gamma_2$ as a kernel is not the same as the product of using the two γ's separately as kernels (even if they are bonafide g's!). A dramatic example is $\gamma_1(x) = x$ and $\gamma_2(x) = \delta(x)$. We have $(x\delta)\{\phi\} \stackrel{\text{def}}{=} \delta\{x\phi\}$ but the latter is zero for all ϕ! (as $x\phi(x) = 0$ at $x = 0$) i.e. we have $(x\delta)$ being *identically zero* as a generalised function even though neither x nor δ are zero as generalised functions.

Fourier series of the delta function

For $f(x) = \delta(x)$ on $-L \leq x \leq L$ we can formally consider a Fourier series (in complex form for convenience)

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{\frac{in\pi x}{L}} \\ \text{with} \quad c_n &= \frac{1}{2L}\int_{-L}^{L} f(x) e^{\frac{-in\pi x}{L}}\, dx. \end{aligned}$$

Thus eq.(2) gives $c_n = \frac{1}{2L}$ for all n and

$$\delta(x) = \frac{1}{2L}\sum_{n=-\infty}^{\infty} e^{\frac{in\pi x}{L}}.$$

Indeed using the RHS expression in $\int_{-L}^{L} \delta(x) f(x)\, dx$ we regain $f(0)$ as the sum of the complex Fourier coefficients c_n of f. Extending periodically to all $\mathbb{R}$ we get

$$\sum_{m=-\infty}^{\infty} \delta(x - 2mL) = \frac{1}{2L}\sum_{n=-\infty}^{\infty} e^{\frac{in\pi x}{L}}. \tag{12}$$

which is sometimes called Poisson's integral formula.

Eigenfunction expansion of the delta function

Revisiting SL theory, let $Y_n(x)$ be an orthonormal family of eigenfunctions for a SL problem on $[a, b]$ with weight function $w(x)$. For $\xi \in (a, b)$, $\delta(x-\xi)$ satisfies homogeneous boundary conditions and we expect to be able to write

$$\delta(x - \xi) = \sum_{n=1}^{\infty} C_n Y_n(x)$$

with

$$C_m = \int_a^b w(x) Y_m(x) \delta(x - \xi)\, dx = w(\xi) Y_m(\xi)$$

so

$$\delta(x - \xi) = w(\xi) \sum_{n=1}^{\infty} Y_n(x) Y_n(\xi).$$

Since $\frac{w(x)}{w(\xi)}\delta(x - \xi) = \delta(x - \xi)$ we can alternatively write

$$\delta(x - \xi) = w(x) \sum_{n=1}^{\infty} Y_n(x) Y_n(\xi). \tag{13}$$

These expressions are consistent with the sampling property since if $g(x) = \sum_{m=1}^{\infty} D_m Y_m(x)$ is the eigenfunction expansion of g then

$$\begin{aligned}\int_a^b g(x)\delta(x-\xi)dx &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} D_m Y_n(\xi)\int_a^b w(x)Y_n(x)Y_m(x)dx,\\ &= \sum_{m=1}^{\infty} D_m Y_m(\xi) = g(\xi),\end{aligned}$$

by orthonormality.

The eigenfunction expansion of the delta function is also intimately related to the eigenfunction expansion of the Green's function that we introduced in §2.7 eq. (**??**) (as we'll see later) and our next task is to develop a theory of Green's functions for solving inhomogeneous ODEs.

7 GREEN'S FUNCTIONS FOR ODEs

Using the concept of a delta function we will now develop a systematic theory of Green's functions for ODEs. Consider a general linear second-order differential operator $\mathcal{L}$ on $[a,b]$, i.e.

$$\mathcal{L}y(x) = \alpha(x)\frac{d^2}{dx^2}y + \beta(x)\frac{d}{dx}y + \gamma(x)y = f(x), \tag{14}$$

where α, β, γ are continuous, $f(x)$ is bounded, and α is nonzero (except perhaps at a finite number of isolated points), and $a \leq x \leq b$ (which may be $\pm\infty$). For this operator $\mathcal{L}$, the **Green's function** $G(x;\xi)$ is defined as the solution to the problem

$$\mathcal{L}G = \delta(x-\xi) \tag{15}$$

satisfying homogeneous boundary conditions $G(a;\xi) = G(b;\xi) = 0$. (Other homogeneous BCs may be entertained too, but for clarity will will treat only the given one here.)

The Green's function has the following fundamental property (established below): the solution to the inhomogeneous problem $\mathcal{L}y = f(x)$ with homogeneous boundary conditions $y(a) = y(b) = 0$ can be expressed as

$$y(x) = \int_a^b G(x;\xi)f(\xi)d\xi. \tag{16}$$

i.e. G is a the kernel of an integral operator that acts as an inverse to the differential operator $\mathcal{L}$. Note that G depends on $\mathcal{L}$, but *not* on the forcing function f, and once G is determined, we are able to construct particular integral solutions for any $f(x)$, directly from the integral formula eq. (16).

We can easily establish the validity of eq. (15) as a simple consequence of (**??**) and the sampling property of the delta function:

$$\begin{aligned}\mathcal{L}\int_a^b G(x;\xi)f(\xi)d\xi &= \int_a^b (\mathcal{L}G)\, f(\xi)\, d\xi, \\ &= \int_a^b \delta(x-\xi)f(\xi)\, d\xi = f(x)\end{aligned}$$

and

$$y(a) = \int_a^b G(a;\xi)f(\xi)\, d\xi = 0 = \int_a^b G(b;\xi)f(\xi)\, d\xi = y(b)$$

since $G(x;\xi) = 0$ at $x = a, b$.

7.1 Construction of the Green's function

We now give a constructive means for determining the Green's function (without recourse to eigenfunctions, although the two approaches must provide equivalent answers). Our construction relies on the fact that for any $x \neq \xi$ away from ξ, $\mathcal{L}G = 0$ so G can be constructed in terms of solutions of the *homogeneous* equation. We proceed with the following steps:

(1) Construct a general homogeneous solution for $x < \xi$ from two linearly independent solutions $y_1(x)$ and $y_2(x)$ to the homogeneous problem $\mathcal{L}y = 0$, and so

$$G(x;\xi) = A(\xi)y_1(x) + B(\xi)y_2(x), \;\; a \leq x < \xi. \tag{17}$$

Here A and B are independent of x, but typically will later be chosen to be dependent on ξ, the upper endpoint of the interval being considered.

(2) Construct a general solution for $x > \xi$ from two linearly independent solutions $Y_1(x)$ and $Y_2(x)$ to the homogeneous problem $\mathcal{L}y = 0$:

$$G(x;\xi) = C(\xi)Y_1(x) + D(\xi)Y_2(x), \;\; \xi < x \leq b. \tag{18}$$

Note: Y_1 and Y_2 will generally *not* be the same as y_1 and y_2 (as other choices will be more convenient, see steps (3) and (4) below), but of course Y_1 and Y_2 must at least be linear combinations of y_1 and y_2.

As for A and B, C and D are independent of x, but typically will depend on ξ. There are thus four constants, (A, B, C, and D) and we need four further conditions to determine G uniquely.

(3) Apply the homogeneous boundary condition at $x = a$ to eliminate either A or B, using the fact that

$$G(a;\xi) = 0 = Ay_1(a) + By_2(a). \tag{19}$$

(4) Similarly, apply the homogeneous boundary condition at $x = b$ to eliminate either C or D, using

$$G(b;\xi) = 0 = CY_1(b) + DY_2(b). \tag{20}$$

Note: we generally choose the forms of y_1, y_2, Y_1, Y_2 to make the application of these two BCs simple.

(5) The two remaining conditions that we need are associated with the properties at the point $x = \xi$. Firstly $G(x, \xi)$ must be **continuous** there, (as we will argue below) and so

$$Ay_1(\xi) + By_2(\xi) = CY_1(\xi) + DY_2(\xi).$$

(6) The fourth and final condition is the **jump** condition, which is the requirement (as we will argue below) that

$$\left[\frac{dG}{dx}\right]_{x=\xi^-}^{x=\xi^+} = \frac{1}{\alpha(\xi)}$$

i.e.

$$\lim_{x\to\xi^+}\frac{dG}{dx} - \lim_{x\to\xi^-}\frac{dG}{dx} = \frac{1}{\alpha(\xi)}$$

and so

$$CY_1'(\xi) + DY_2'(\xi) - Ay_1'(\xi) - By_2'(\xi) = \frac{1}{\alpha(\xi)},$$

where $\alpha(x)$ is the coefficient of the second derivative in the operator $\mathcal{L}$ as defined in eq. (14).

Note: often, *but not always*, the operator $\mathcal{L}$ has $\alpha(x) = 1$. Also everything here assumes that the right hand side of eq. (15) is $+1 \times \delta(x - \xi)$ whereas some applications have an extra factor $K\delta(x - \xi)$ which must first be scaled out before applying our formulae from here.

(7) With the four conditions in (3), (4), (5) and (6), the Green's function is uniquely determined, and then for any given forcing function $f(x)$, the solution to the forced problem with homogeneous boundary conditions $y(a) = 0 = y(b)$ is

$$\begin{aligned} y(x) &= \int_a^b G(x;\xi) f(\xi) d\xi \\ &= Y_1(x) \int_a^x C(\xi) f(\xi) d\xi + Y_2(x) \int_a^x D(\xi) f(\xi) d\xi \\ &+ \ y_1(x) \int_x^b A(\xi) f(\xi) d\xi + y_2(x) \int_x^b B(\xi) f(\xi) d\xi. \end{aligned}$$

Note that the ξ integral $\int_a^b$ is separated at x into two parts (i) $\int_a^x$ and (ii) $\int_x^b$. In the range of (i) we have $\xi < x$ so the formula eq. (18) for $G(x;\xi)$ with coefficients $C(\xi), D(\xi)$ is applicable even though this expression incorporated the BC at $x = b$. For (ii) we have $x > \xi$ so we use the $G(x;\xi)$ expression from eq. (17), that incorporated the BC at $x = a$.

The conditions on $G(x;\xi)$ at $x = \xi$

We are constructing $G(x;\xi)$ from pieces of homogeneous solutions (as required by eq. (15) for all $x \neq \xi$) which are restrictions of well behaved (differentiable) functions on (a, b). At $x = \xi$ eq. (15) imposes further conditions on how the two homogeneous pieces must match up.

The continuity condition: first suppose there was a jump discontinuity at $x = \xi$. Then $\frac{d}{dx}G \propto \delta(x - \xi)$ and $\frac{d^2}{dx^2}G \propto \delta'(x - \xi)$. However eq. (15) shows that $\frac{d^2}{dx^2}G$ cannot involve a generalised function of the form δ' (but only δ). Hence G cannot have a jump discontinuity at $x = \xi$ and it must be continuous there. Note however that it is fine for G' to have a jump discontinuity at $x = \xi$, and indeed we have:

The jump condition: eq. (15) imposes a constraint on the size of any prospective jump discontinuity in dG/dx at $x = \xi$, as follows – integrating eq. (15) over an arbitrarily small interval about $x = \xi$ gives:

$$\int_{\xi-\epsilon}^{\xi+\epsilon} \alpha(x) \left[\frac{d^2G}{dx^2}\right] dx + \int_{\xi-\epsilon}^{\xi+\epsilon} \beta(x) \left[\frac{dG}{dx}\right] dx + \int_{\xi-\epsilon}^{\xi+\epsilon} \gamma(x) G \, dx = \int_{\xi-\epsilon}^{\xi+\epsilon} \delta(x - \xi) dx = 1.$$

Writing the three LHS integrals as

$$T_1 + T_2 + T_3 = 1.$$

we have
$T_3 \to 0$ as $\epsilon \to 0$ (since G is continuous and γ is bounded);
$T_2 \to 0$ as $\epsilon \to 0$ (as dG/dx and β are bounded).
Also α is continuous and $\alpha(x) \to \alpha(\xi)$ as $\epsilon \to 0$ so

$$T_1 \to \alpha(\xi) \lim_{\epsilon \to 0} \int_{\xi-\epsilon}^{\xi+\epsilon} \left[\frac{d^2G}{dx^2}\right] dx = \alpha(\xi) \left[\frac{dG}{dx}\right]_{x=\xi^-}^{x=\xi^+},$$

giving the jump condition.

Example of construction of a Green's function

Consider the problem

$$\mathcal{L}y = -y'' - y = f(x) \quad y(0) = y(1) = 0.$$

We follow our algorithm:

1. For $0 \le x < \xi$, $G'' + G = 0$, which suggests $G = A\cos x + B \sin x$ (especially noting the $x = 0$ BC).

2. For $\xi < x \le 1$, $G'' + G = 0$, which suggests (cf $x = 1$ BC!)

$$G = C\cos(1-x) + D\sin(1-x).$$

   These expressions are motivated by looking at the BCs. G here could of course be expressed in terms of the basic $\sin x$ and $\cos x$ homogeneous solutions i.e.

$$\begin{aligned} G &= (C\cos 1 + D\sin 1)\cos x + (C\sin 1 - D\cos 1)\sin x, \\ &= \hat{C}\cos x + \hat{D}\sin x. \end{aligned}$$

   but then see step 4 below!

3. Applying the $x = 0$ boundary condition $G(0, \xi) = 0$ gives $A = 0$.

4. Applying the $x = 1$ boundary condition $G(1, \xi) = 0$ gives simply $C = 0$. Note that for the $\hat{C}, \hat{D}$ form we would have got $\hat{C} = -\hat{D}\tan 1$ and things just get more messy in steps 5 and 6 below.

5. Therefore,

$$G(x;\xi) = \begin{cases} B\sin x & 0 < x < \xi; \\ D\sin(1-x) & \xi < x < 1. \end{cases}$$

   Applying the continuity condition,

$$B = D\frac{\sin(1-\xi)}{\sin\xi}.$$

6. Therefore,

$$G(x;\xi) \;=\; \begin{cases} D\frac{\sin(1-\xi)\sin x}{\sin\xi} & 0<x<\xi; \\ D\sin(1-x) & \xi<x<1. \end{cases}$$

In our operator $\mathcal{L}$ we have $\alpha(x)=-1$ and so the jump condition is that

$$D\left[-\cos(1-x)\right]_{\xi^+} - D\left[\frac{\sin(1-\xi)\cos x}{\sin\xi}\right]_{\xi^-} \;=\; -1$$

and we get

$$D=\frac{\sin\xi}{\sin 1},$$

which then means that the Green's function is:

$$G(x;\xi) \;=\; \begin{cases} \frac{\sin(1-\xi)\sin x}{\sin 1} & 0<x<\xi; \\ \frac{\sin(1-x)\sin\xi}{\sin 1} & \xi<x<1. \end{cases}$$

7. And so we are able to construct the complete solution to $-y''-y=f(x)$ as (taking care to use the second G formula in the first integral, and the first G formula in the second integral!):

$$y(x)=\frac{\sin(1-x)}{\sin 1}\int_0^x f(\xi)\sin\xi d\xi+\frac{\sin x}{\sin 1}\int_x^1 f(\xi)\sin(1-\xi)d\xi. \qquad (21)$$

General solution for self adjoint $\mathcal{L}$s

If $\mathcal{L}$ is self adjoint i.e. $\beta(x)=d\alpha/dx$ (and we have homogeneous BCs $y(a)=y(b)=0$) then we can give a general formula for the Green's function. Let y_1 and y_2 be two independent solutions of the homogeneous equation $\mathcal{L}y=0$ which satisfy the BCs at $x=a$ and $x=b$ *respectively* (so not the same as y_1, y_2 used above!) Then consider

$$G(x;\xi)=\frac{y_1(x)y_2(\xi)H(\xi-x)+y_2(x)y_1(\xi)H(x-\xi)}{J(y_1,y_2)}$$

where

$$J(y_1,y_2)=\alpha(x)W(y_1,y_2)=\alpha(x)[y_1(x)y_2'(x)-y_2(x)y_1'(x)]$$

is called the *conjunct* (and W is the Wronskian that you met last year). Since $\mathcal{L}$ is self adjoint $J(y_1,y_2)$ is necessarily a (non-zero) constant (exercise – show this e.g. consider dJ/dx and use $\mathcal{L}y_1=\mathcal{L}y_2=0$ so $y_1\mathcal{L}y_2-y_2\mathcal{L}y_1=0$).

Note that the $H(\xi-x)$ and $H(x-\xi)$ factors above imply that for $x<\xi$ (resp. $x>\xi$) only the first (resp. the second) terms appear. Now it is straightforward to verify that the $G(x;\xi)$ formula above satisfies the continuity and jump conditions at $x=\xi$ (exercise – e.g. consider $x\to\xi$ from above and below separately, and for the jump condition remember that $J(y_1,y_2)$ is a constant). Thus the above formula provides a neat way of immediately writing down the Green's function.

This formula also shows that the Green's function is **symmetric** in variables x and ξ i.e. $G(x;\xi) = G(\xi;x)$ as featured in our example above, and also seen in our previous expression (in § 2.7) of $G(x;\xi)$ in terms of a series of eigenfunctions Y_n viz.

$$G(x;\xi) = \sum_{n=1}^{\infty} \frac{Y_n(x)Y_n(\xi)}{\lambda_n}.$$

Green's functions for inhomogeneous BCs

Homogeneous BCs were essential to the notion of a Green's function (since in eq. (16) the integral represents a kind of "continuous superposition" of solutions for individual ξ values). However we can also treat problems with *inhomogeneous* BCs using a standard trick to reduce them to homogeneous BC problems:
(1) First find a particular solution y_p to the *homogeneous equation* $\mathcal{L}y = 0$ satisfying the *in*homogeneous BCs (usually easy).
(2) Then using the Green's function we solve

$$\mathcal{L}y_g = f \quad \text{with } \textit{homogeneous} \text{ BCs} \quad y(a) = y(b) = 0.$$

(3) Then since $\mathcal{L}$ is linear the solution of $\mathcal{L}y = f$ with the inhomogeneous BCs is given by $y = y_p + y_g$.

Example. Consider $-y'' - y = f(x)$ with inhomogeneous BCs $y(0) = 0$ and $y(1) = 1$.
The general solution of the homogeneous equation $-y'' - y = 0$ is $c_1 \cos x + c_2 \sin x$ and the (inhomogeneous) BCs require $c_1 = 0$ and $c_2 = 1/\sin 1$ so $y_p = \sin x/\sin 1$.
Using the Green's function for this $\mathcal{L}$ calculated in the previous example, we can write down y_g (solution of $\mathcal{L}y = f$ with homogeneous BCs) as given in eq. (21) and so the final solution is

$$y(x) = \frac{\sin x}{\sin 1} + \frac{\sin(1-x)}{\sin 1} \int_0^x f(\xi) \sin \xi d\xi + \frac{\sin x}{\sin 1} \int_x^1 f(\xi) \sin(1-\xi) d\xi.$$

Equivalence of eigenfunction expansion of $G(x;\xi)$

For self adjoint $\mathcal{L}$s (with homogeneous BCs) we have two different expressions for the Green's function

$$G(x;\xi) = \frac{y_1(x)y_2(\xi)H(\xi - x) + y_2(x)y_1(\xi)H(x-\xi)}{J(y_1, y_2)} \tag{22}$$

and

$$G(x;\xi) = \sum_{n=1}^{\infty} \frac{Y_n(x)Y_n(\xi)}{\lambda_n}. \tag{23}$$

The first is in terms of *homogeneous solutions* y_1 and y_2 ($\mathcal{L}y = 0$) whereas the second is in terms of *eigenfunctions* Y_n and eigenvalues λ_n of the corresponding SL system $\mathcal{L}Y_n = \lambda_n Y_n w(x)$.

In § 2.7 we derived eq. (23) without any mention of delta functions, but it may also be quickly derived using the eigenfunction expansion of $\delta(x-\xi)$ given in eq. (13) viz.

$$\delta(x-\xi) = w(x)\sum_{n=1}^{\infty} Y_n(x)Y_n(\xi).$$

Indeed viewing ξ as a parameter and writing the Green's function as an eigenfunction expansion

$$G(x;\xi) = \sum_{n=1}^{\infty} B_n(\xi)Y_n(x) \tag{24}$$

then $\mathcal{L}G = \delta(x-\xi)$ gives

$$\begin{aligned}\mathcal{L}G &= \sum_{n=1}^{\infty} B_n(\xi)\mathcal{L}Y_n(x) = \sum_{n=1}^{\infty} B_n(\xi)\lambda_n w(x)Y_n(x)\\ &= \delta(x-\xi) = w(x)\sum_{n=1}^{\infty} Y_n(x)Y_n(\xi).\end{aligned}$$

Multiplying the two end terms by $Y_m(x)$ and integrating from a to b we get (by the weighted orthogonality of the Y_m's)

$$B_m(\xi)\lambda_m = Y_m(\xi)$$

and then eq. (24) gives the expression eq. (23).

Remark. Note that the formula eq. (23) requires that all eigenvalues λ be *nonzero* i.e. that the homogeneous equation $\mathcal{L}y = 0$ (also being the eigenfunction equation for $\lambda = 0$) should have no nontrivial solutions satisfying the BCs. Indeed the existence of such solutions is problematic for the concept of a Green's function, as providing an inverse operator for $\mathcal{L}$: if nontrivial y_0 with $\mathcal{L}y_0 = 0$ exist, then the inhomogeneous equation $\mathcal{L}y = f$ will not have a unique solution (since if y is any solution then so is $y + y_0$) and then the operator $\mathcal{L}$ is not invertible. This is just the infinite dimensional analogue of the familiar situation of a system of linear equations $A\underline{x} = \underline{b}$ with non-invertible coefficient matrix A (and indeed a matrix is non-invertible iff it has nontrivial eigenvectors belonging to eigenvalue zero). $\square$

Example. As an illustrative example of the equivalence, consider again $\mathcal{L}y = -y'' - y$ on $[a,b] = [0,1]$ with BCs $y(0) = y(1) = 0$.

The normalised eigenfunctions and corresponding eigenvalues are easily calculated to be

$$Y_n(x) = \sqrt{2}\sin n\pi x \qquad \lambda_n = n^2\pi^2 - 1$$

so by eq. (23) we have

$$G(x;\xi) = 2\sum_{n=1}^{\infty} \frac{\sin n\pi x\,\sin n\pi\xi}{n^2\pi^2 - 1}. \tag{25}$$

On the other hand we had in a previous example the expression constructed from homogeneous solutions:

$$G(x;\xi) = \frac{\sin(1-x)\,\sin\xi\,H(x-\xi) + \sin x\,\sin(1-\xi)\,H(\xi-x)}{\sin 1}$$

and using trigonometric addition formulae we get

$$G(x;\xi) = \cos x\,\sin\xi\,H(x-\xi) + \sin x\,\cos\xi\,H(\xi-x) - \cot 1\,\sin x\,\sin\xi\,. \qquad (26)$$

Now comparing eqs (25,26) and viewing ξ as a parameter and x as the independent variable, we see that the equivalence of the two expressions amounts to the Fourier sine series in eq. (25) of the function (for each ξ) in eq. (26)

$$\begin{aligned} f(x) &= \cos x\,\sin\xi\,H(x-\xi) + \sin x\,\cos\xi\,H(\xi-x) - \cot 1\,\sin x\,\sin\xi \\ &= \textstyle\sum_{n=1}^{\infty} b_n(\xi)\sin n\pi x. \end{aligned}$$

The Fourier coefficients are given as usual by (noting that $\sin n\pi x$ has norm $1/\sqrt{2}$):

$$b_n(\xi) = 2\int_0^1 f(x)\sin n\pi x\,dx$$

and a direct (but rather tedious) calculation (exercise?...) gives

$$b_n(\xi) = \frac{2\sin n\pi\xi}{n^2\pi^2 - 1}$$

as expected. (In this calculation note that the Heaviside functions $H(x-\xi)$ and $H(\xi-x)$ merely alter the integral limits from $\int_0^1$ to $\int_\xi^1$ and $\int_0^\xi$ respectively).

7.2 Physical interpretation of the Green's function

We can think of the expression

$$y(x) = \int_a^b G(x;\xi)\,f(\xi)\,d\xi$$

as a 'summation' (integral) of individual point source effects, of strength $f(\xi)$ with $G(x;\xi)$ characterising the elementary effect (as a function of x) of a unit point source placed at ξ.

To illustrate with a physical example consider again the wave equation for a horizontal elastic string with gravity and ends fixed at $x = 0, L$. If $y(x,t)$ is the (small) transverse displacement we had

$$T\frac{\partial^2 y}{\partial x^2} - \mu g = \mu\frac{\partial^2 y}{\partial t^2} \qquad 0 \le x \le L \qquad y(0) = y(L) = 0.$$

Here T is the (constant) tension in the string and μ is the mass density per unit length which may be a function of x. We consider the steady state solution $\partial y/\partial t = 0$ i.e. $y(x)$ is the shape of a (non-uniform) string hanging under gravity, and

$$\frac{d^2y}{dx^2} = \frac{\mu(x)g}{T} \equiv f(x) \tag{27}$$

is a forced self adjoint equation (albeit having a very simple $\mathcal{L}y = y''$).

Case 1 (massive uniform string): if μ is constant eq. (27) is easily integrated and setting $y(0) = y(L) = 0$ we get the parabolic shape

$$y = \frac{\mu g}{2T}x(x - 2L).$$

Case 2 (light string with point mass at $x = \xi$): consider a point mass concentrated at $x = \xi$ i.e. the mass density is $\mu(x) = m\delta(x - \xi)$ with $\mu = 0$ for $x \neq \xi$. For $x \neq \xi$ the string is massless so the only force acting is the (tangential) tension so the string must be straight either side of the point mass. To find the location of the point mass let θ_1 and θ_2 be the angles either side. Then resolving forces vertically the equilibrium condition is

$$mg = T(\sin\theta_1 + \sin\theta_2) \approx T(\tan\theta_1 + \tan\theta_2)$$

where we have used the small angle approximation $\sin\theta \approx \tan\theta$ (and y hanging, is negative). Thus

$$y(\xi) = \frac{mg}{T}\frac{\xi(\xi - L)}{L}$$

for the point mass at $x = \xi$. Hence from *physical principles* we have derived the shape for a point mass at ξ:

$$y(x) = \frac{mg}{T} \times \begin{cases} \frac{x(\xi-L)}{L} & 0 \leq x \leq \xi; \\ \frac{\xi(x-L)}{L} & \xi \leq x \leq L; \end{cases} \tag{28}$$

which is the solution of eq. (27) with forcing function $f(x) = \frac{mg}{T}\delta(x - \xi)$.

Next let's calculate the Green's function for eq. (27) i.e. solve

$$\mathcal{L}G = \frac{d^2G}{dx^2} = \delta(x - \xi) \quad \text{with } G(0;\xi) = G(L;\xi) = 0.$$

Summarising our algorithmic procedure we have successively:
(1) For $0 < x < \xi$, $G = Ax + B$.
(2) For $\xi < x < L$, $G = C(x - L) + D$.
(3) BC at zero implies $B = 0$.
(4) BC at L implies $D = 0$.
(5) Continuity implies $C = A\xi/(\xi - L)$.
(6) The jump condition gives $A = \frac{(\xi-L)}{L}$ so finally

$$G = \begin{cases} \frac{x(\xi-L)}{L} & 0 \leq x \leq \xi; \\ \frac{\xi(x-L)}{L} & \xi \leq x \leq L; \end{cases}$$

and we see that eq. (28) is precisely the Green's function with a multiplicative scale for a point source of strength mg/T.

Case 3 (continuum generalisation): if in case 2 we had several point masses m_k at $x_k = \xi_k$ we can sum the solutions to get

$$y(x) = \sum_k \frac{m_k\, g}{T} G(x; \xi_k). \tag{29}$$

For the continuum limit we imagine $(N-1)$ masses $m_k = \mu(\xi_k)\Delta\xi$ placed at equal intervals $x_k = \xi_k = k\Delta\xi$ with $\Delta\xi = L/N$ and $k = 1, \ldots, N-1$. Then by the Riemann sum definition of integrals, as $N \to \infty$, eq. (29) becomes

$$y(x) = \int_0^L \frac{g\,\mu(\xi)}{T} G(x; \xi)\, d\xi.$$

If μ is constant this function reproduces the parabolic result of case 1 as you can check by direct integration (exercise, taking care with the limits of integration).

7.3 Application of Green's functions to IVPs

Green's functions can also be used to solve initial value problems. Consider the problem (viewing now the independent variable as time t)

$$\mathcal{L}y = f(t), \; t \geq a, \; y(a) = y'(a) = 0. \tag{30}$$

The algorithm for construction of the Green's function is very similar to the previous BVP method. As before, we want to find G such that $\mathcal{L}G = \delta(t - \tau)$.

1. Construct G for $a \leq t < \tau$ as a general solution of the homogeneous equation: $G = Ay_1(t) + By_2(t)$.

2. But now, apply **both** boundary (i.e. initial) conditions to this solution:

   $$\begin{aligned} Ay_1(a) + By_2(a) &= 0, \\ Ay_1'(a) + By_2'(a) &= 0. \end{aligned}$$

   Since y_1 and y_2 are linearly independent the determinant of the coefficient matrix (being the Wronskian) is non-zero and we get $A = B = 0$, and so $G(t; \tau) = 0$ for $a \leq t < \tau$!

3. Construct G for $\tau \leq t$, again as a general solution of the homogeneous equation: $G = Cy_1(t) + Dy_2(t)$.

4. Now apply the continuity and jump conditions at $t = \tau$ (noting that $G = 0$ for $t < \tau$). We thus get

   $$\begin{aligned} Cy_1(\tau) + Dy_2(\tau) &= 0, \\ Cy_1'(\tau) + Dy_2'(\tau) &= \frac{1}{\alpha(\tau)}, \end{aligned}$$

where $\alpha(t)$ is as usual the coefficient of the second derivative in the differential operator $\mathcal{L}$. This determines $C(\tau)$ and $D(\tau)$ completing the construction of the Green's function $G(t;\tau)$.

For forcing function $f(t)$ we thus have

$$y(t) = \int_a^b f(\tau)G(t;\tau)\,d\tau$$

but since $G = 0$ for $\tau > t$ (by step 2 above) we get

$$y(t) = \int_a^t f(\tau)G(t;\tau)\,d\tau$$

i.e. the solution at time t depends only on the input (forcing) for **earlier** times $a \leq \tau \leq t$. Physically this is a causality condition.

Example of an IVP Green's function

Consider the problem

$$\frac{d^2y}{dt^2} + y \quad = \quad f(t), \;\; y(0) = y'(0) = 0. \tag{31}$$

Following our procedure above we get (exercise)

$$G(t;\tau) \quad = \quad \begin{cases} 0 & 0 \leq t \leq \tau; \\ C\cos(t-\tau) + D\sin(t-\tau) & t \geq \tau. \end{cases}$$

Note our judicious choice of independent solutions for $t \geq \tau$: continuity at $t = \tau$ implies simply that $C = 0$, while the jump condition ($\alpha(\tau) = 1$) implies that $D = 1$, and so

$$G(t;\tau) \quad = \quad \begin{cases} 0 & 0 \leq t \leq \tau; \\ \sin(t-\tau) & t \geq \tau, \end{cases}$$

which gives the solution as

$$y(t) = \int_0^t f(\tau)\sin(t-\tau)d\tau. \tag{32}$$

7.4 Higher order differential operators

We mention that there is a natural generalization of Green's functions to higher order differential operators (and indeed PDEs, as we shall see in the last part of the course). If $\mathcal{L}y = f(x)$ is a n^{th}-order ODE (with the coefficient of the highest derivative one for simplicity, and $n > 2$) with homogeneous boundary conditions on $[a,b]$, then

$$y(x) = \int_a^b f(\xi)G(x;\xi)d\xi,$$

where

- G satisfies the homogeneous boundary conditions;
- $\mathcal{L}G = \delta(x - \xi)$;
- G and its first $n - 2$ derivatives continuous at $x = \xi$;
- $d^{(n-1)}G/dx^{(n-1)}(\xi^+) - d^{(n-1)}G/dx^{(n-1)}(\xi^-) = 1$.

See example sheet III for an example.

8 FOURIER TRANSFORMS

8.1 From Fourier series to Fourier transforms

Fourier series provide a very useful tool for working with periodic functions (or functions on a finite domain $[0, T]$ which may then be extended periodically) and here we seek to develop this facility to apply to non-periodic functions, on the infinite domain $\mathbb{R}$. We begin by imagining allowing the period to tend to infinity, in the Fourier series formalism.

Suppose f has period T and Fourier series (in complex form)

$$f(t) = \sum_{r=-\infty}^{\infty} c_r e^{iw_r t} \tag{33}$$

where $w_r = \frac{2\pi r}{T}$. We can write $w_r = r\Delta w$ with frequency gap $\Delta w = 2\pi/T$. The coefficients are given by

$$c_r = \frac{1}{T}\int_{-T/2}^{T/2} f(u)e^{-iw_r u}\,du = \frac{\Delta w}{2\pi}\int_{-T/2}^{T/2} f(u)e^{-iw_r u}\,du. \tag{34}$$

For this integral to exist in the limit as $T \to \infty$ we'll require that $\int_{-\infty}^{\infty} |f(x)|\,dx$ exists, but then the $1/T$ factor implies that for each r, $c_r \to 0$ too. Nevertheless we can substitute eq. (34) into eq. (33) to get

$$f(t) = \sum_{r=-\infty}^{\infty} \frac{\Delta w}{2\pi} e^{iw_r t}\int_{-T/2}^{T/2} f(u)e^{-iw_r u}\,du \tag{35}$$

Now recall the Riemann sum definition of the integral of a function g:

$$\sum_{r=-\infty}^{\infty} \Delta w\, g(w_r) \to \int_{-\infty}^{\infty} g(w)\,dw$$

with w becoming a continuous variable. For eq. (35) we take

$$g(w_r) = \frac{e^{iw_r t}}{2\pi}\int_{-T/2}^{T/2} f(u)e^{-iw_r u}\,du$$

(with t viewed as a parameter) and letting $T \to \infty$ we get

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dw\, e^{iwt}\left[\int_{-\infty}^{\infty} f(u)e^{-iwu}\,du\right]. \tag{36}$$

Introduce the **Fourier transform** (FT) of f, denoted $\tilde{f}$:

$$\tilde{f}(w) = \int_{-\infty}^{\infty} f(t)e^{-iwt}\,dt \tag{37}$$

and then eq. (36) becomes

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(w)\, e^{iwt}\, dw \tag{38}$$

which is **Fourier's inversion theorem**.

Thus (glossing over issues of when the integrals converge) the FT is a mapping from functions $f(t)$ to functions $g(w) = \tilde{f}(w)$ where we conventionally use variable names t and w respectively. But we can write g as a function of t and consider *its* FT $\tilde{g}$. Note that there is a close structural similarity between the FT formula and its inverse eqs. (37,38) – interchanging the roles of t and $-w$ and correcting the 2π factor. This leads to the following useful dual relation:

$$\text{if } g(x) = \tilde{f}(x) \text{ then } \tilde{g}(w) = 2\pi f(-w). \tag{39}$$

Thus we can immediately write down the FT of any function $g(x)$ if it is known to be the FT of some other function f.

Warning: different books give slightly different definitions of FT. Generally we can have $\tilde{f}(w) = A \int_{-\infty}^{\infty} f(t) e^{\pm iwt}\, dt$ and then the inversion formula is

$$f(t) = \frac{1}{2\pi A} \int_{-\infty}^{\infty} \tilde{f}(w) e^{\mp iwt}\, dt$$

(pairing opposite $\pm$ signs). Quite common is $A = \frac{1}{\sqrt{2\pi}}$, symmetrically giving the same constant for the FT and inverse FT formulas, but in this course we will always use $A = 1$ as above. Some other books also use $e^{\pm 2\pi iwt}$ in the integrals.

Remarks.
• The dual pair of variables t and w above are referred to by different names in different applications. If t is time, w is frequency; if t is space x then w is often written as k (wavenumber) or p (momentum, especially in quantum mechanics) and k-space may be called spectral space.
• the Fourier transform is an example of an *integral transform* with *kernel* e^{-iwt}. There are other important integral transforms such as the Laplace transform with kernel e^{-st} (s real, and integration range 0 to ∞) that you'll probably meet later.
• if f has a finite jump discontinuity at t then (just as for Fourier series) the inversion formula eq. (38) returns the average value $\frac{f(t_+)+f(t_-)}{2}$.

8.2 Properties of the Fourier transform

Let f and g have FTs $\tilde{f}$ and $\tilde{g}$ respectively. Then the following properties follow easily from the integral expressions eqs (37,38). (Here λ and μ are constants).
• **(Linearity)** $\lambda f + \mu g$ has FT $\lambda \tilde{f} + \mu \tilde{g}$;
• **(Translation)** if $g(x) = f(x - \lambda)$ then $\tilde{g}(k) = e^{-ik\lambda} \tilde{f}(k)$;
• **(Frequency shift)** if $g(x) = e^{i\lambda x} f(x)$ then $\tilde{g}(k) = \tilde{f}(k - \lambda)$. Note: this also follows from the translation property and the duality in eq. (39).

• **(Scaling)** if $g(x) = f(\lambda x)$ then $\tilde{g}(k) = \tilde{f}(k/\lambda)$;
• **(Multiplication by x)** if $g(x) = xf(x)$ then $\tilde{g}(k) = i\tilde{f}'(k)$ (applying integration by parts in the FT integral for g).
• **(FT of a derivative)** dual to the multiplication rule (or applying integration by parts to $\int_{-\infty}^{\infty} f'(x)e^{-ikx}\,dx$) we have the derivative rule: if $g(x) = f'(x)$ then $\tilde{g}(k) = ik\tilde{f}(k)$ i.e. differentiation in physical space becomes multiplication in frequency space (as we had for Fourier series previously).

The last property can be very useful for solving (linear) differential equations in physical space – taking the FT of the equation can lead to a simpler equation in k-space (algebraic equation for an ODE, or an ODE from a PDE if we FT on one or more variables). Thus we solve the simpler problem and invert the answer back to physical space. The last step can involve difficult inverse-FT integrals, and in some important applications techniques of complex methods/analysis are very effective.

Example. (Dirichlet's discontinuous formula)

Consider the "top hat" (English) or "boxcar" (American) function

$$f(x) \;=\; \begin{cases} 1 & |x| \le a, \\ 0 & |x| > a \end{cases} \tag{40}$$

for $a > 0$. Its FT is easy to calculate:

$$\tilde{f}(k) = \int_{-a}^{a} e^{-ikx}\,dx = \int_{-a}^{a} \cos(kx)\,dx = \frac{2\sin(ka)}{k}. \tag{41}$$

so by the Fourier inversion theorem we get

$$\frac{1}{\pi}\int_{-\infty}^{\infty} e^{ikx}\frac{\sin ka}{k}\,dk = \begin{cases} 1 & |x| < a, \\ 0 & |x| > a. \end{cases} \tag{42}$$

Now setting $x = 0$ (so the integrand is even, and rewriting the variable k as x) we get **Dirichlet's discontinuous formula**:

$$\int_0^\infty \frac{\sin(ax)}{x}\,dx \;=\; \begin{cases} \frac{\pi}{2} & a > 0, \\ 0 & a = 0, \\ -\frac{\pi}{2} & a < 0, \end{cases} \tag{43}$$

$$\;=\; \frac{\pi}{2}\mathrm{sgn}(a), \tag{44}$$

(To get the $a < 0$ case we've simply used the fact that $\sin(-ax) = -\sin(ax)$.)

The above integral (forming the basis of the inverse FT of $\tilde{f}$) is quite tricky to do without the above FT inversion relations. It can be done easily using the elegant methods of complex contour integration (Cauchy's integral formula) but a direct "elementary" method is the following: introduce the two parameter(!) integral

$$I(\lambda, \alpha) = \int_0^\infty \frac{\sin(\lambda x)}{x}e^{-\alpha x}\,dx \tag{45}$$

with $\alpha > 0$ and λ real. Then

$$\begin{aligned}\frac{\partial I}{\partial \lambda} &= \int_0^\infty \cos(\lambda x) e^{-\alpha x} dx = \Re\left(\int_0^\infty e^{-(\alpha+\lambda i)x} dx\right), \\ &= \Re\left[\frac{-e^{-(\alpha+\lambda i)x}}{\alpha + \lambda i}\right]_0^\infty = \Re\left(\frac{1}{\alpha+\lambda i}\right) = \frac{\alpha}{\alpha^2+\lambda^2}.\end{aligned}$$

Now integration w.r.t. λ gives

$$I(\lambda,\alpha) = \arctan\left(\frac{\lambda}{\alpha}\right) + C(\alpha)$$

with integration constant $C(\alpha)$ for each α. But from eq. (45) $I(0,\alpha) = 0$ so $C(\alpha) = 0$. Then considering $I(\lambda, 0) = \lim_{\alpha\to 0} I(\lambda,\alpha)$ we have $I(\lambda,0) = \lim_{\alpha\to 0} \arctan(\lambda/\alpha)$ which gives eq. (43).

8.3 Convolution and Parseval's theorem for FTs

In applications it is often required to take the inverse FT of a product of FTs i.e. we want to find $h(x)$ such that $\tilde{h}(k) = \tilde{f}(k)\tilde{g}(k)$ where $\tilde{f}$ and $\tilde{g}$ are FTs of known functions f and g respectively. Applying the definitions of the Fourier transform and its inverse, we can write

$$\begin{aligned} h(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{f}(k)\tilde{g}(k) e^{ikx} dk, \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{g}(k) e^{ikx}\left[\int_{-\infty}^{\infty} f(u) e^{-iku} du\right] dk. \end{aligned}$$

Now (assuming the f and $\tilde{g}$ are absolutely integrable) we can change the order of integration to write

$$\begin{aligned} h(x) &= \int_{-\infty}^{\infty} f(u)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{g}(k) e^{ik(x-u)} dk\right] du, \\ h(x) &= \int_{-\infty}^{\infty} f(u) g(x-u)\, du = f * g, \end{aligned} \tag{46}$$

This integrated combination $h = f*g$ is called the **convolution** of f and g. Convolution is called *faltung* (or 'folding') in German, which picturesquely describes the way in which the functions are combined – the graph of g is flipped ("folded") about the variable vertical line $u = x/2$ and then integrated against f.

By exploiting the dual structure of eq. (39) of FT and inverse FT, the above result readily shows (exercise) that a product of functions in physical (x) space has a FT that is the convolution of the individual FTs (with a 2π factor):

$$h(x) = f(x)g(x) \to \tilde{h}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{f}(u)\tilde{g}(k-u) du. \tag{47}$$

Parseval's theorem for FTs

An important simple corollary of (46) is Parseval's theorem for FTs. Let $g(x) = f^*(-x)$. Then

$$\begin{aligned}\tilde{g}(k) &= \int_{-\infty}^{\infty} f^*(-x)e^{-ikx}\,dx \\ &= \left[\int_{-\infty}^{\infty} f(-x)e^{ikx}\,dx\right]^* = \left[\int_{-\infty}^{\infty} f(y)e^{-iky}\,dy\right]^* = \tilde{f}^*(k).\end{aligned}$$

Thus by eq. (46) the inverse FT of $\tilde{f}(k)\tilde{f}^*(k)$ is the convolution of $f(x)$ and $f^*(-x)$ i.e.

$$\int_{-\infty}^{\infty} f(u)f^*(u-x)du = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{f}(k)|^2 e^{ikx}dk.$$

Setting $x = 0$ gives the **Parseval formula for FTs**:

$$\int_{-\infty}^{\infty} |f(u)|^2\,du = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{f}(k)|^2\,dk. \tag{48}$$

Thus the FT as a mapping from functions $f(x)$ to functions $\tilde{f}(k)$ preserves the squared norm of the function (up to a constant $1/2\pi$ factor).

8.4 The delta function and FTs

There is a natural extension of the notion of FT to generalised functions. Here we will indicate some basic features relating to the Dirac delta function, using intuitive arguments based on formally manipulating integrals (without worrying too much about whether they converge or not). The results can be rigorously justified using the Schwarz function formalism for generalised functions that we outlined in chapter 7, and we will indicate later how this formalism embraces FTs too (optional section below).

Writing $f(x)$ as the inverse of its FT we have

$$\begin{aligned}f(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\left[\int_{-\infty}^{\infty} f(u)e^{-iku}du\right]dk, \\ &= \int_{-\infty}^{\infty} f(u)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-u)}dk\right]du.\end{aligned}$$

Comparing this with the sampling property of the delta function, we see that the term in the square brackets must be a representation of the delta function:

$$\delta(u-x) = \delta(x-u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-u)}\,dk. \tag{49}$$

In particulkar setting $u = 0$ we get the FT pair

$$f(x) = \delta(x) \leftrightarrow \tilde{f}(k) = 1. \tag{50}$$

Note that this is also consistent with directly putting $f(x) = \delta(x)$ into the basic FT formula $\tilde{f}(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx$ and using the sampling property of δ.

Eq. (50) can be used to obtain further FT pairs by formally applying basic FT rules to it. The dual relation eq. (39) gives

$$f(x) = 1 \quad \leftrightarrow \quad \tilde{f}(k) = \int_{-\infty}^{\infty} e^{-ikx}\,dx = 2\pi\delta(k).$$

From the translation property of FTs (or applying the sampling property of $\delta(x-a)$) we get,

$$f(x) = \delta(x-a) \quad \leftrightarrow \quad \tilde{f}(k) = \int_{-\infty}^{\infty} \delta(x-a)e^{-ikx}dx = e^{-ika}$$

and the dual relation

$$f(x) = e^{iax} \quad \leftrightarrow \quad \tilde{f}(k) = 2\pi\delta(k-a).$$

Then $e^{\pm i\theta} = \cos\theta \pm i\sin\theta$ gives

$$\begin{aligned} f(x) = \cos(\omega x) \quad &\leftrightarrow \quad \tilde{f}(k) = \pi\left[\delta(k+\omega) + \delta(k-\omega)\right], \\ f(x) = \sin(\omega x) \quad &\leftrightarrow \quad \tilde{f}(k) = \pi i\left[\delta(k+\omega) - \delta(k-\omega)\right] \end{aligned} \tag{51}$$

So, a highly localised signal in physical space (e.g. a delta function) has a very spread out representation in spectral space. Conversely a highly spread out (yet periodic) signal in physical space is highly localised in spectral space. This is illustrated too in exercise sheet 3, in computing the FT of a Gaussian bell curve.

Towards a rigorous theory of FTs of generalised functions

(optional section)

Many of the FT integrals above actually don't technically converge! (e.g. FT(1) which we identified as $2\pi\delta(k)$). Here we outline how to make rigorous sense of the above results, extending our previous discussion of generalied functions in terms of Schwarz functions.

If f is any suitably regular ordinary function on $\mathbb{R}$ and ϕ is any Schwarz function then it is easy to see from the FT definition eq. (37) that

$$\int_{-\infty}^{\infty} \tilde{f}(x)\phi(x)\,dx = \int_{-\infty}^{\infty} f(x)\tilde{\phi}(x)\,dx.$$

Here as usual the tilde denotes FT and we have written the independent variables as x rather than t or k. Hence if F and $\tilde{F}$ denote the functionals on $\mathcal{S}$ (space of Schwarz functions) associated to f and $\tilde{f}$ then we have

$$\tilde{F}\{\phi\} = F\{\tilde{\phi}\}$$

and we now **define**(!) the FT of *any* generalised function by this formula. For example for the delta function we get

$$\tilde{\delta}\{\phi\} \overset{(a)}{=} \delta\{\tilde{\phi}\} \overset{(b)}{=} \tilde{\phi}(0) \overset{(c)}{=} \int_{-\infty}^{\infty} 1.\phi(t)\,dt.$$

(where (a) is by definition, (b) is the action of the delta function and (c) is by setting $w = 0$ in $\tilde{\phi}(w) = \int_{-\infty}^{\infty} \phi(t)e^{-iwt}\,dt$). Thus comparing to the formal notation of generalised function kernels

$$\tilde{\delta}\{\phi\} \equiv \int_{-\infty}^{\infty} \tilde{\delta}(x)\phi(x)\,dx$$

we have $\tilde{\delta}(x) = 1$ as expected. Similarly all our formulas above may be rigorously established.

Example (FT of the Heaviside function H).

$$H(x) = \begin{cases} 1 & x > 0 \\ \frac{1}{2} & x = 0 \\ 0 & x < 0 \end{cases}$$

We begin by noting that $H(x) = \frac{1}{2}(\mathrm{sgn}(x) + 1)$ and recalling Dirichlet's discontinuous formula

$$\int_{-\infty}^{\infty} \frac{\sin kx}{k}\,dk = 2\int_0^{\infty} \frac{\sin kx}{k}\,dk = \pi \mathrm{sgn}(x).$$

Thus for the FT of $\mathrm{sgn}(x)$ we have

$$\begin{aligned}
\widetilde{\mathrm{sgn}}\{\phi\} &= \int_{-\infty}^{\infty} \mathrm{sgn}(s)\,\tilde{\phi}(s)\,ds \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{\phi}(s)\frac{\sin us}{\pi u}\,du\,ds \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{\phi}(s)\frac{e^{ius} - e^{-ius}}{2\pi i u}\,du\,ds \\
&= \int_{-\infty}^{\infty} \frac{1}{iu}\left[\phi(u) - \phi(-u)\right] du \\
&= 2\int_{-\infty}^{\infty} \left[\frac{1}{ix}\right] \phi(x)\,dx.
\end{aligned}$$

(In the second last line we have used the Fourier inversion formula, and to get the last line we have substituted $u = x$ and $u = -x$ respectively in the two terms of the integral). Thus comparing with

$$\widetilde{\mathrm{sgn}}\{\phi\} = \int_{-\infty}^{\infty} \widetilde{\mathrm{sgn}}(k)\phi(k)\,dk$$

we get

$$\widetilde{\mathrm{sgn}}(k) = \frac{2}{ik}.$$

Finally (using $\mathrm{FT}(1) = 2\pi\delta(k)$) we get

$$\tilde{H}(k) = \mathrm{FT}\left(\frac{1}{2}(\mathrm{sgn}(x) + 1)\right) = \frac{1}{ik} + \pi\delta(k).$$

It is interesting and instructive here to recall that $H'(x) = \delta(x)$ so from the differentiation rule for FTs we get

$$ik\tilde{H}(k) = \tilde{\delta}(k) = 1$$

even though $\tilde{H}(k)$ is *not* $\frac{1}{ik}$ but equals $\frac{1}{ik} + \pi\delta(k)$! However this is *not inconsistent* since using the correct formula for $\tilde{H}(k)$, we have

$$ik\tilde{H}(k) = ik(\frac{1}{ik} + \pi\delta(k)) = 1 + i\pi k\delta(k)$$

and as noted previously (top of page 67) $x\delta(x)$ is actually the zero generalised function, so we can correctly deduce that $ik\tilde{H}(k) = 1$!

8.5 FTs and linear systems, transfer functions

FTs are often used in the systematic analysis of linear systems which arise in many engineering applications. Suppose we have a linear operator $\mathcal{L}$ acting on input $\mathcal{I}(t)$ to give output $\mathcal{O}(t)$ e.g. we may have an amplifier that can in general change the amplitude and phase of a signal.

Using FTs we can express the physical input signal via an inverse FT as

$$\mathcal{I}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{\mathcal{I}}(w)e^{iwt}\,dw$$

which is called the **synthesis** of the input, expressing it as a combination of components with various frequencies each having amplitude and phase given by the modulus and argument respectively, of $\tilde{\mathcal{I}}(w)$. The FT itself is known as the **resolution** of the pulse (into its frequency components):

$$\tilde{\mathcal{I}}(w) = \int_{-\infty}^{\infty} \mathcal{I}(t)e^{-iwt}\,dt.$$

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

$$\mathcal{O}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{R}(w)\tilde{\mathcal{I}}(w)e^{iwt}\,dw. \qquad (52)$$

$\tilde{R}(w)$ is called the **transfer function** of the system and its inverse FT $R(t)$ is called the **response function**.

Warning: in various texts both $R(t)$ and $\tilde{R}(w)$ are referred to as the response or transfer function, in different contexts.

Thus

$$\tilde{R}(w) = \int_{-\infty}^{\infty} R(t)e^{-iwt}\,dt \qquad R(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{R}(w)e^{iwt}\,dw$$

and looking at eq. (52) we see that $R(t)$ is the output $\mathcal{O}(t)$ of the system when the input has $\tilde{\mathcal{I}}(w) = 1$ i.e. when the input is $\delta(t)$ i.e. $\mathcal{L}R(t) = \delta(t)$. Hence $R(t)$ is closely related to the Green's function when $\mathcal{L}$ is a differential operator.

Eq. (52) also shows that $\mathcal{O}(t)$ is the inverse FT of a product $\tilde{R}(w)\tilde{\mathcal{I}}(w)$ so the convolution theorem gives

$$\mathcal{O}(t) = \int_{-\infty}^{\infty} \mathcal{I}(u)R(t-u)\,du. \qquad (53)$$

Next consider the important extra physical condition of **causality**. Assume that there is no input signal before $t = 0$ i.e. $\mathcal{I}(t) = 0$ for $t < 0$, and suppose the system has zero output for zero input (e.g. "the amplifier does not hum..") so $R(t) = 0$ for $t < 0$. Then in eq, (53) the lower limit can be set to zero (as $\mathcal{I}(u) = 0$ for $u < 0$) and the upper limit to t (as $R(t-u) = 0$ for $u > t$):

$$\mathcal{O}(t) = \int_0^t \mathcal{I}(u)R(t-u)\,du \tag{54}$$

which is formally the same as our previous expressions in chapter 7.3 with Green's functions for IVPs.

General form of transfer functions for ODEs

In many applications the relationship between input and output is given by a linear finite order ODE:

$$\mathcal{L}_m[\mathcal{I}(t)] = \left(\sum_{j=0}^{m} b_{m-j}\frac{d^j}{dt^j}\right)[\mathcal{I}(t)] = \left(\sum_{i=0}^{n} a_{n-i}\frac{d^i}{dt^i}\right)[\mathcal{O}(t)] = \mathcal{L}_n[\mathcal{O}(t)]. \tag{55}$$

For simplicity here we will consider the case $m = 0$ so the input acts directly as a forcing term. Taking FTs we get

$$\begin{aligned}\tilde{\mathcal{I}}(\omega) &= \left[a_n + a_{n-1}i\omega + \ldots + a_1(i\omega)^{n-1} + a_0(i\omega)^n\right]\tilde{\mathcal{O}}(\omega)\\ \tilde{R}(\omega) &= \frac{1}{\left[a_n + a_{n-1}i\omega + \ldots + a_1(i\omega)^{n-1} + a_0(i\omega)^n\right]}.\end{aligned}$$

So the transfer function is a rational function with an n^{th} degree polynomial as the denominator.

The denominator of $\tilde{R}(w)$ can be factorized into a product of roots of the form $(i\omega - c_j)^{k_j}$ for $j = 1, \ldots, J$ (allowing for repeated roots) where $k_j \geq 1$ and $\sum_{j=1}^{J} k_j = n$). Thus

$$\tilde{R} = \frac{1}{(i\omega - c_1)^{k_1}\ldots(i\omega - c_J)^{k_J}}$$

and using partial fractions, this can be expressed as a simple sum of terms of the form

$$\frac{\Gamma_{mj}}{(i\omega - c_j)^m},\ 1 \leq m \leq k_j$$

where Γ_{mj} are constants.

So we have

$$\tilde{R}(\omega) = \sum_j\sum_m \frac{\Gamma_{mj}}{(i\omega - c_j)^m}$$

and to find the response function $R(t)$ we need to Fourier-invert a function of the form

$$\tilde{h_m}(\omega) = \frac{1}{(i\omega - \alpha)^{m+1}},\ m \geq 0.$$

We give the answer and check (by Fourier-transforming it) that it works.

Consider the function $h_0(t) = e^{\alpha t}$ for $t > 0$, and zero otherwise i.e. $h_0(t) = 0$ for $y < 0$. Therefore

$$\begin{aligned}\tilde{h_0}(\omega) &= \int_0^\infty e^{(\alpha - i\omega)t} dt \\ &= \left[\frac{e^{(\alpha-i\omega)t}}{\alpha - i\omega}\right]_0^\infty \\ &= \frac{1}{i\omega - \alpha}\end{aligned}$$

provided $\Re(\alpha) < 0$. So $h_0(t)$ is identified.

Remark: If $\Re(\alpha) > 0$ then the above integral is divergent. Indeed recalling the theory of linear constant coefficient ODEs, we see that the c_j's above are the roots of the auxiliary equation and the equation has solutions with terms $e^{c_j t}$ which grow exponentially if $\Re(c_j) > 0$. Such exponentially growing functions are problematic for FT and inverse FT integrals so here we will consider only the case $\Re(c_j) < 0$ for all j, corresponding to 'stable' ODEs whose solutions do not grow unboundedly as $t \to \infty$.

Next consider the function $h_1(t) = te^{\alpha t}$ for $t > 0$ and zero otherwise, so $h_1(t) = th_0(t)$. Recalling the "multiplication by x" rule for FTs viz:

$$\text{if } f(x) \to \tilde{f}(w) \text{ then } xf(x) \to i\tilde{f}'(w),$$

we get

$$\tilde{h}_1(w) = i\frac{d}{dw}\left(\frac{1}{iw - \alpha}\right) = \frac{1}{(iw - \alpha)^2}$$

(which may also be derived directly by evaluating the FT integral $\int_0^\infty te^{\alpha t}e^{-iwt}\,dt$).

Similarly (or proof by induction) we have (for $Re(\alpha) < 0$)

$$h_m(t) = \begin{cases} \frac{t^m e^{\alpha t}}{m!} & t > 0; \\ 0 & t \leq 0 \end{cases} \quad \leftrightarrow \quad \tilde{h_m}(\omega) = \frac{1}{(i\omega - \alpha)^{m+1}}, \; m \geq 0$$

and so it is possible to construct the output from the input using eq. (54) for such stable systems easily. Physically, see that functions of the form $h_m(t)$ always decay as $t \to \infty$, but they can increase initially to some finite time maximum (at time $t_m = m/|\alpha|$ if $\alpha < 0$ and real for example). We also see that $R(\xi) = 0$ for $\xi < 0$ so in eq. (53) the upper integration limit can be replaced by t, so $\mathcal{O}(t)$ depends only on the input at earlier times, as expected.

Example: the forced damped oscillator

The relationship between Green's functions and response functions can be nicely illustrated by considering the linear operator for a damped oscillator:

$$\frac{d^2}{dt^2}y + 2p\frac{d}{dt}y + (p^2 + q^2)y = f(t), \quad p > 0. \tag{56}$$

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 $y(0) = y'(0) = 0$. Taking FTs we get

$$\begin{aligned} (i\omega)^2\tilde{y} + 2ip\omega\tilde{y} + (p^2+q^2)\tilde{y} &= \tilde{f} \\ \text{so} \quad \tilde{R}\tilde{f} = \frac{\tilde{f}}{-\omega^2 + 2ip\omega + (p^2+q^2)} &= \tilde{y} \end{aligned}$$

and taking inverse FTs we get

$$\begin{aligned} y(t) &= \int_0^t R(t-\tau)f(\tau)d\tau, \\ \text{with} \quad R(t-\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega(t-\tau)}}{p^2+q^2+2ip\omega-\omega^2}d\omega. \end{aligned}$$

Now consider $\mathcal{L}R(t-\tau)$, using this integral formulation, and assuming that formal differentiation within the integral sign is valid

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

using (49). Therefore, the Green's function $G(t;\tau)$ **is** the response function $R(t-\tau)$ by (mutual) definition. On sheet 3, question 1 you are asked to fill in the details, computing both $R(t)$ and the Green's function explicitly.

Example. FTs can also be used for ODE problems on the full line $\mathbb{R}$ so long as the functions involved are required to have suitably good asymptotic properties for the FT integrals to exist. As an example consider

$$\frac{d^2y}{dx^2} - A^2y = -f(x) \qquad -\infty < x < \infty$$

such that $y \to 0$, $y' \to 0$ as $|x| \to \infty$, and A is a positive real constant. Taking FTs we have

$$\tilde{y} = \frac{\tilde{f}}{A^2+k^2}, \tag{57}$$

and we seek to identify $g(x)$ such that

$$\tilde{g} = \frac{1}{A^2+k^2}.$$

Consider

$$h(x) = \frac{e^{-\mu|x|}}{2\mu}, \ \mu > 0.$$

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

$$\begin{aligned}\tilde{h}(k) &= \Re\left(\frac{1}{\mu}\int_0^\infty \exp\left[-x(\mu+ik)\right]dx\right)\\ &= \frac{1}{\mu}\Re\left(\frac{1}{\mu+ik}\right) = \frac{1}{\mu^2+k^2}\end{aligned}$$

so we have identified $g(x) = \frac{e^{-A|x|}}{2A}$. The convolution theorem then gives

$$y(x) = \frac{1}{2A}\int_{-\infty}^{\infty} f(u)\exp(-A|x-u|)du. \tag{58}$$

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 BCs on the Green's function for $|x|\to\infty$).

8.6 FTs and discrete signal processing

Another hugely important application of FTs is to the theory and practice of discrete signal processing i.e. the manipulation and analysis of data that is sampled at discrete times, as occurring for example in any digital representation of music or images (CDs, DVDs) and all computer graphics and sound files etc. Here we will give just a very brief overview of a few fundamental ingredients of this subject, which could easily justify an entire course on its own.

Consider a signal $h(t)$ which is sampled at evenly spaced time intervals Δ apart:

$$h_n = h(n\Delta) \qquad n = \ldots -2,-1,0,1,2,\ldots$$

The **sampling rate** or **sampling frequency** is $f_s = 1/\Delta$ (samples per unit time) with angular frequency $w_s = 2\pi f_s$. Consider two complex exponential signals with pure frequencies

$$h_1(t) = e^{iw_1 t} = e^{2\pi i f_1 t} \qquad h_2(t) = e^{iw_2 t} = e^{2\pi i f_2 t}.$$

These will have precisely the same Δ samples when $(f_1-f_2)\Delta$ is an integer. We can avoid this possibility by choosing Δ so that $|f_1-f_2|\Delta<1$. A general signal $h(t)$ is called $w_{\max}$-**bandwidth limited** if its FT $\tilde{h}(w)$ is zero for $|w|>w_{\max}$ i.e. $\tilde{h}$ is supported entirely in $[-w_{\max},w_{\max}]$. Writing $f_{\max}=w_{\max}/2\pi$ we see that we can distinguish different $f_{\max}$-bandwidth limited signals with Δ-sampling if

$$2f_{\max}\Delta < 1 \quad \text{i.e.} \quad \Delta < \frac{1}{2f_{\max}} \tag{59}$$

which is called the **Nyquist condition** (with Nyquist frequency $f = 1/2\Delta$ for sampling interval Δ). For a given sampling interval Δ violation of eq. (59) leads to the phenomenon of **aliasing** i.e. an inability to distinguish properties of signals with frequencies exceeding half the sampling rate. Via the relation $(f_1 - f_2)\Delta =$ integer, such high frequencies f_1 will be "aliased" onto lower frequencies f_2 within $[-f_{\max}, f_{\max}]$.

Reconstruction of a signal from its samples – the sampling theorem

Suppose g is a $w_{\max}$-bandwidth limited signal and write $f_{\max} = w_{\max}/2\pi$. From the definition of the Fourier transform we have

$$\tilde{g}(w) = \int_{-\infty}^{\infty} g(t)e^{-iwt}dt,$$

and the inversion formula with the bandwidth limit gives

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{g}(w)e^{iwt}dw = \frac{1}{2\pi}\int_{-w_{\max}}^{w_{\max}} \tilde{g}(w)e^{iwt}\,dw.$$

Now let us set the sampling interval to be $\Delta = 1/(2f_{\max}) = \pi/(w_{\max})$. so samples are taken at $t_n = n\Delta = n\pi/w_{\max}$ giving

$$g(t_n) = g_n \quad = \quad \frac{1}{2\pi}\int_{-w_{\max}}^{w_{\max}} \tilde{g}(w)\exp\left[\frac{i\pi n w}{w_{\max}}\right]\,dw.$$

Now recalling our formulas for complex Fourier *series* coefficients of a function on $[a,b] = [-w_{\max}, w_{\max}]$ we see that the g_{-n} are $w_{\max}/\pi$ times the complex Fourier coefficients c_n for a function $\tilde{g}_p(w)$ which coincides with $\tilde{g}(w)$ on $(-w_{\max}, w_{\max})$ and is extended to all $\mathbb{R}$ as a periodic function with period $2w_{\max}$ i.e.

$$\tilde{g}_p(w) = \frac{\pi}{w_{\max}}\sum_{n=-\infty}^{\infty} g_n \exp\left(\frac{-in\pi w}{w_{\max}}\right).$$

Therefore the actual (bandwidth limited) Fourier transform of the original function is the product of the periodically repeating $\tilde{g}_p(\omega)$ and a box-car function $\tilde{h}(w)$ as defined in (40):

$$\begin{aligned}
\tilde{h}(w) &= \begin{cases} 1 & |w| \le w_{\max}, \\ 0 & \text{otherwise.} \end{cases}, \\
\tilde{g}(w) &= \tilde{g}_p(w)\tilde{h}(w) \\
&= \left[\frac{\pi}{w_{\max}}\sum_{n=-\infty}^{\infty} g_n \exp\left(\frac{-in\pi w}{w_{\max}}\right)\right]\tilde{h}(w),
\end{aligned}$$

This is an *exact* equality: the countably infinite discrete sequence of samples $g(n\Delta) = g_n$ of a bandwidth-limited function $g(t)$, completely determines its Fourier transform $\tilde{g}(w)$.

We can now apply the Fourier inversion formula to *exactly* reconstruct the *full* signal $g(t)$ for all t, from its samples at $t = n\Delta$. Assuming that swapping the order of integration and summation is fine, we have

$$g(t) \quad = \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{g}(w)e^{iwt}dw,$$

$$
\begin{aligned}
&= \frac{1}{2w_{\max}} \sum_{n=-\infty}^{\infty} g_n \int_{-w_{\max}}^{w_{\max}} \exp\left(iw \left[t - \frac{n\pi}{w_{\max}} \right] \right) dw, \\
&= \frac{1}{2w_{\max}} \sum_{n=-\infty}^{\infty} g_n \left[\frac{\exp(i[w_{\max}t - \pi n]) - \exp(-i[w_{\max}t - \pi n])}{i\left(t - \frac{n\pi}{w_{\max}}\right)} \right] \\
&= \sum_{n=-\infty}^{\infty} g_n \frac{\sin(w_{\max}t - \pi n)}{w_{\max}t - \pi n}, \\
&= \sum_{n=-\infty}^{\infty} g(n\Delta) \operatorname{sinc}[w_{\max}(t - n\Delta)]
\end{aligned}
$$

where $\operatorname{sinc}(t) = \frac{\sin t}{t}$.

In summary, the bandwidth-limited function can be represented (for continuous time) exactly by this representation in terms of its discretely sampled values, with the continuous values being filled in by this expression which is known as the Shannon-Whittaker sampling formula. This result is called the **(Shannon) sampling theorem**. Such full reconstruction of bandwidth-limited functions, by multiplying by the 'Whittaker sinc' function or cardinal function $\operatorname{sinc}(t)$ centred on the sampling points and summing, is at the heart of all digital music reproduction.

8.7 The discrete Fourier transform

For any natural number N the **discrete Fourier transform** (mod N) denoted DFT is defined to be the N by N matrix with entries

$$
[\mathrm{DFT}]_{mn} = e^{-\frac{2\pi i}{N}mn} \qquad m, n = 0, 1, 2, \ldots, (N-1). \tag{60}
$$

Note that here we number rows and columns from 0 to $N-1$ rather than the more conventional 1 to N. Thus the first row and column are always all 1's and DFT is a symmetric matrix. Some books include a $1/\sqrt{N}$ pre-factor (just like a $1/\sqrt{2\pi}$ factor that we did not use in our definition of FT).

In further pure mathematical theory (group representation theory) there is a general notion of a Fourier transform on any group G. It may be shown that the Fourier transform of functions on $\mathbb{R}$, the Fourier series of periodic finctions on $\mathbb{R}$, and DFT, are all examples of this single construction, just using different groups viz. respectively $\mathbb{R}$ with addition, the circle group $\{e^{i\theta} : 0 \le \theta < 2\pi\}$ with multiplication, and the group $\mathbb{Z}_N$ of integers mod N with addition. Correspondingly DFT enjoys a variety of nice properties analogous to those we've seen for Fourier series and transforms. We now derive some of these from the definition of DFT above. So far we have that DFT is a linear mapping from $\mathbb{C}^N$ to $\mathbb{C}^N$.

The inverse DFT

$\frac{1}{\sqrt{N}}$DFT is actually a *unitary* matrix i.e. the inverse is the adjoint (conjugate transpose):

$$\left[\frac{1}{\sqrt{N}}\mathrm{DFT}\right]^{-1} = \left[\frac{1}{\sqrt{N}}\mathrm{DFT}\right]^{\dagger} \quad \text{so} \quad \left[\frac{1}{\sqrt{N}}\mathrm{DFT}\right]^{\dagger}\left[\frac{1}{\sqrt{N}}\mathrm{DFT}\right] = I$$

so

$$\mathrm{DFT}^{-1} = \frac{1}{N}\mathrm{DFT}^{\dagger}.$$

To see this write $w = e^{-2\pi i/N}$ and note that each row or column of DFT is a geometric sequence $1, r, r^2, \ldots, r^{N-1}$ with ratio r being a power of w (depending on the choice of row or column). Now recall the basic property of such series of N^{th} roots of unity: if $r = w^a$ for any $a \neq 0$ (or more generally a is not a multiple of N) then $r \neq 1$ but $r^N = 1$ so

$$1 + r + r^2 + \ldots + r^{N-1} = \frac{1 - r^N}{1 - r} = 0$$

but if $a = 0$ (or is a multiple of N) then $r = 1$ so

$$1 + r + r^2 + \ldots + r^{N-1} = 1 + 1 + \ldots + 1 = N.$$

Using this, it is easy to see (exercise) that the set of rows (or set of columns) of $\mathrm{DFT}/\sqrt{N}$ is an *orthonormal set* of vectors in $\mathbb{C}^N$ and hence that $(\frac{1}{N}\mathrm{DFT}^{\dagger})(\mathrm{DFT}) = I$.

A Parseval equality (also known as **Plancherel's theorem**)

For any $\underline{a} = (a_0, \ldots, a_{N-1}) \in \mathbb{C}^N$, if $\underline{\tilde{a}} = (\tilde{a}_0, \ldots, \tilde{a}_{N-1}) = \mathrm{DFT}\,\underline{a}$ then

$$\underline{a} \cdot \underline{a} = \frac{1}{N}\underline{\tilde{a}} \cdot \underline{\tilde{a}}\,.$$

This follows immediately from the general relationship of adjoints to inner products viz. $(A\underline{u}, \underline{v}) = (\underline{u}, A^{\dagger}\underline{v})$. Hence

$$(\underline{\tilde{a}}\,,\, \underline{\tilde{a}}) = (\mathrm{DFT}\,\underline{a}\,,\, \mathrm{DFT}\,\underline{a}) = (\underline{a}\,,\, \mathrm{DFT}^{\dagger}\,\mathrm{DFT}\,\underline{a}) = N(\underline{a}\,,\, \underline{a}).$$

Convolution property

For any $\underline{a} = (a_0, \ldots, a_{N-1})$ and $\underline{b} = (b_0, \ldots, b_{N-1})$ introduce the **convolution** $\underline{c} = \underline{a} * \underline{b}$ defined by

$$c_k = \sum_{m=0}^{N-1} a_m b_{k-m} \qquad k = 0, \ldots, N-1$$

(where the subscript $k - m$ is computed mod N to be in $0, \ldots, N-1$). Then DFT maps convolutions into straightforward component-wise products:

$$\tilde{c}_k = \tilde{a}_k\,\tilde{b}_k \qquad \text{for each } k = 0, \ldots, N-1\ .$$

To see this, use the DFT and convolution definitions to write

$$\tilde{c}_k = \sum_{l=0}^{N-1} c_l\, w^{kl} = \sum_{m=0}^{N-1}\sum_{l=0}^{N-1} a_m\, b_{l-m}\, w^{kl}$$

and then substitute for index l introducing $p = l - m$ to get $\tilde{a}_k \tilde{b}_k$ directly on RHS (exercise).

Remark (FFT) (Optional).

DFT has a wide variety of applications in both pure and applied mathematics. One of its most important properties is the existence of a so-called **fast Fourier transform** (FFT) which refers to the following idea: suppose we want to compute DFT $\underline{a}$ for some $\underline{a} \in \mathbb{C}^N$ and N is large. Direct matrix multiplication will require $O(N^2)$ basic arithmetic operations (additions and multiplications) – for each of the N components of DFT $\underline{a}$ we need $O(N)$ operations to compute the inner product of a row of DFT and the column vector $\underline{a}$. Now because of the very special structure of the DFT matrix entries (as a pattern of roots of unity) it can be shown that if $N = 2^M$ is a power of 2, then there is a faster algorithm to compute DFT $\underline{a}$, requiring only $O(N \log N)$ basic arithmetic operations! – a dramatic exponential ($N \to \log N$) saving in run time. This is the so-called FFT algorithm, often attributed to a paper by Cooley and Tukey from 1965, but it already appears in the notebooks of Gauss from 1805 – he invented it on the side, as an aid to speed up his by-hand calculations of trajectories of asteroids!

An important application of FFT is to provide a "super-fast" algorithm for integer multiplication. If $a = a_{n-1} \dots a_0$ and $b = b_{n-1} \dots b_0$ are two n digit numbers (written as usual in terms of decimal or binary digits) then the direct calculation of the product $c = ab$ by long multiplication takes $O(n^2)$ elementary arithmetic steps (additions and multiplications of single digits). However if we write $a = \sum_{m=0}^{n-1} a_m 10^m$ and similarly for b, then the coefficient of 10^k (for any k) in the product ab is the *convolution* $\sum_{m=0}^{n-1} a_m b_{k-m}$. With this observation we can construct a multiplication algorithm along the following lines: view a and b as vectors of their digits; take DFTs ($O(n \log n)$ steps with FFT) to get $\tilde{a}, \tilde{b}$; form the entry-wise product $\tilde{c}$ of $\tilde{a}$ and $\tilde{b}$ ($O(n)$ steps); finally take inverse DFT of $\tilde{c}$ to get the digits of c ($O(n \log n)$ steps again). Thus we have $O(n \log n)$ steps in all compared to $O(n^2)$ for standard long multiplication. For example (assuming all the big-O constants to be 1) if the numbers have 1000 digits, then we get (only!) about $n \log n = 3000$ steps compared to about $n^2 = 1$ million steps!

Example (finite sampling)

The sampling theorem showed how in some circumstances (viz. bandwidth limited signals) we can reconstruct a signal $h(t)$ (and hence also its FT) exactly from a discrete but infinite set of samples $h(n\Delta)$. But realistically in practice we can obtain only a finite number N of samples say $h_m = h(t_m)$ for $t_m = m\Delta$ and $m = 0, 1, \dots, N-1$, and hence expect only an approximation to its FT $\tilde{h}$, at some points. The DFT arises in the computation of such approximations.

Suppose $h(t)$ is negligible outside $[0, T]$ and we sample N points as above with $\Delta = T/N$. Now consider the FT $\tilde{h}$ at w with frequency $f = w/2\pi$. By approximating the FT integral as a Riemann sum using our sampled values h_m in $[0, T]$, we have

$$\tilde{h}(w) = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t}\, dt \tag{61}$$

$$\approx \quad \Delta \sum_{m=0}^{N-1} h_m e^{-2\pi i f(m\Delta)}. \tag{62}$$

Since FT is an invertible map, from our data of N samples of $h(t)$ we'd like to estimate N samples $\tilde{h}(w_n)$ of the FT at frequencies $f_n = w_n/2\pi$. We know that if $|f|$ exceeds the Nyquist frequency $f_c = 1/2\Delta$ associated to the sampling rate Δ, then features with such frequencies become indistinguishably aliased into frequencies with $|f| < f_c$. Hence we'll choose our N discrete frequency values f_n to be equally spaced in $[-f_c, f_c]$ i.e.

$$f_n = \frac{n}{N\Delta} \qquad n = -N/2, \ldots, 0, \ldots, N/2 - 1.$$

Subsitiuting these into the above approximation we get

$$\tilde{h}(f_n) = \Delta \sum_{m=0}^{N-1} h_m e^{-\frac{2\pi i}{N} mn}$$

so if we introduce vectors $\underline{\tilde{h}}$ and $\underline{h}$ whose components are the discrete values of $\tilde{h}$ and h, then the above gives

$$\underline{\tilde{h}} = \Delta \, \mathrm{DFT}(\underline{h}).$$

Remark (quantum computing) (optional)
DFT (or more precisely the unitary matrix $\mathrm{QFT} = \mathrm{DFT}/\sqrt{N}$) has a fundamental significance in *quantum mechanics*. In quantum theory, states of physical systems are described by complex vectors (finite dimensional for many physical properties) and physical time evolution is represented by a *unitary* operation on the vectors. Thus QFT represents an allowable quantum physical state transformation and further to FFT (as a mathematical transformation of a list of N complex numbers), it can be shown that QFT as a *quantum physical transformation* on a physical quantum state can be physically implemented in $O((\log N)^2)$ elementary quantum steps i.e. exponentially faster than the $O(N \log N)$-time FFT on classical data. This is a very remarkable property with momentous consequences: if we had a "quantum computer" that could implement elementary quantum operations on quantum states as its computational steps (rather than the conventional Boolean operations on classical bit values) then the superduper-fast QFT could be exploited (as a consequence of a little bit of number theory...) to provide an algorithm for integer factorisation that can factorise any integer of n digits in $O(n^3)$ steps, whereas the best known classical factoring algorithm has a profoundly slower run time of about $O(e^{n^{1/3}})$ steps. Because of this time speed up, such a quantum computer would be able to efficiently (i.e. realistically in actual practice) crack currently used cryptosystems (e.g. RSA) whose security relies on the hardness of factoring, and decipher much encrypted secret information that exists the public domain. Because all known classical factoring algorithms are so profoundly slower, such decryption is *in practice* impossible on a regular computer - it can in principle be done but would typically need an estimated running time well exceeding the age of the universe to finish the task. The tripos part III course titled "Quantum computation" provides a detailed exposition of QFT and the amazing quantum factoring algorithm.