

Mathematics 412

Partial Differential Equations

© S. A. Fulling

Fall, 2005

## The Wave Equation

This introductory example will have three parts.\*

1. I will show how a particular, simple partial differential equation (PDE) arises in a physical problem.
2. We'll look at its solutions, which happen to be unusually easy to find in this case.
3. We'll solve the equation again by *separation of variables*, the central theme of this course, and see how *Fourier series* arise.

The *wave equation* in two variables (one space, one time) is

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},$$

where  $c$  is a constant, which turns out to be the speed of the waves described by the equation.

Most textbooks derive the wave equation for a vibrating string (e.g., Haberman, Chap. 4). It arises in many other contexts — for example, light waves (the electromagnetic field). For variety, I shall look at the case of sound waves (motion in a gas).

### SOUND WAVES

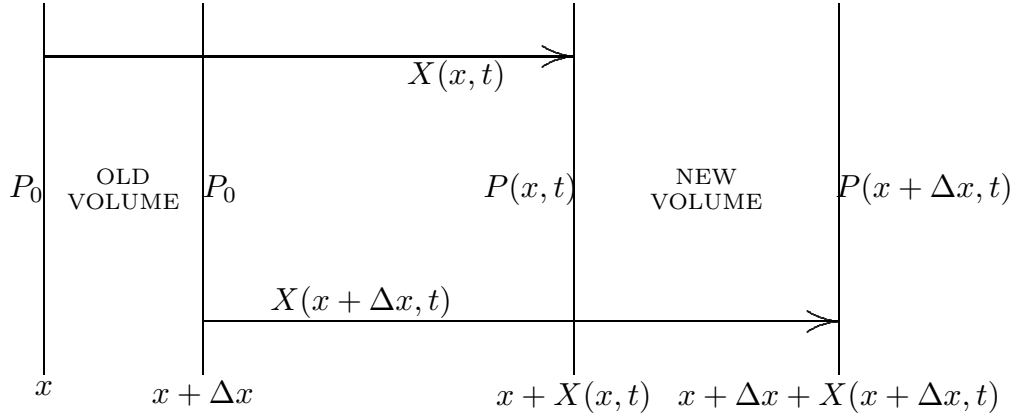
**Reference:** *Feynman Lectures in Physics*, Vol. 1, Chap. 47.

We assume that the gas moves back and forth in one dimension only (the  $x$  direction). If there is no sound, then each bit of gas is at rest at some place  $(x, y, z)$ . There is a uniform equilibrium density  $\rho_0$  (mass per unit volume) and pressure  $P_0$  (force per unit area). Now suppose the gas moves; all gas in the layer at  $x$  moves the same distance,  $X(x)$ , but gas in other layers move by different distances. More precisely, at each time  $t$  the layer originally at  $x$  is displaced to  $x + X(x, t)$ . There it experiences a new density and pressure, called

$$\rho = \rho_0 + \rho_1(x, t), \quad P = P_0 + P_1(x, t).$$

---

\* Simultaneously, students should be reading about another introductory example, the *heat equation*, in Chapters 1 and 2 of Haberman's book. (See also Appendix A of these notes.)



Given this scenario, Newton's laws imply a PDE governing the motion of the gas. The input to the argument is three physical principles, which will be translated into three equations that will imply the wave equation.

**I. The motion of the gas changes the density.** Take a slab of thickness  $\Delta x$  in the gas at rest. The total amount of gas in the slab (measured by mass) is

$$\rho_0 \times \text{volume} = \rho_0 \Delta x \times \text{area}.$$

We can consider a patch with area equal to 1. In the *moving* gas at time  $t$ , this *same* gas finds itself in a new volume (area times thickness)

$$(\text{area} \times) \{[x + \Delta x + X(x + \Delta x, t)] - [x + X(x, t)]\} \equiv \Delta x_{\text{new}}.$$

(Cancel  $x$ .) Thus  $\rho_0 \Delta x = \rho \Delta x_{\text{new}}$ . If  $\Delta x$  is small, we have

$$X(x + \Delta x, t) - X(x, t) \approx \frac{\partial X}{\partial x} \cdot \Delta x;$$

$$\rho_0 \Delta x = \rho \left( \Delta x + \frac{\partial X}{\partial x} \Delta x \right).$$

(Cancel  $\Delta x$ .) So

$$\rho_0 = (\rho_0 + \rho_1) \frac{\partial X}{\partial x} + \rho_0 + \rho_1.$$

Since  $\rho_1 \ll \rho_0$ , we can replace  $\rho_0 + \rho_1$  by  $\rho_0$  in its first occurrence — but not the second, where the  $\rho_0$  is cancelled, leaving  $\rho_1$  as the most important term. Therefore, we have arrived (essentially by geometry) at

$$\rho_1 = -\rho_0 \frac{\partial X}{\partial x}. \tag{I}$$

II. **The change in density corresponds to a change in pressure.** (If you push on a gas, it pushes back, as we know from feeling balloons.) Therefore,  $P = f(\rho)$ , where  $f$  is some *increasing* function.

$$P_0 + P_1 = f(\rho_0 + \rho_1) \approx f(\rho_0) + \rho_1 f'(\rho_0)$$

since  $\rho_1$  is small. (Cancel  $P_0$ .) Now  $f'(\rho_0)$  is greater than 0; call it  $c^2$ :

$$P_1 = c^2 \rho_1. \tag{II}$$

III. **Pressure inequalities generate gas motion.** The force on our slab (measured positive to the right) equals the pressure acting on the left side of the slab minus the pressure acting on the right side (times the area, which we set to 1). But this force is equal to mass times acceleration, or

$$(\rho_0 \Delta x) \frac{\partial^2 X}{\partial t^2}.$$

$$\rho_0 \Delta x \frac{\partial^2 X}{\partial t^2} = P(x, t) - P(x + \Delta x, t) \approx - \frac{\partial P}{\partial x} \Delta x.$$

(Cancel  $\Delta x$ .) But  $\partial P_0 / \partial x = 0$ . So

$$\rho_0 \frac{\partial^2 X}{\partial t^2} = - \frac{\partial P_1}{\partial x}. \tag{III}$$

Now put the three equations together. Substituting (I) into (II) yields

$$P_1 = -c^2 \rho_0 \frac{\partial X}{\partial x}.$$

Put that into (III):

$$\rho_0 \frac{\partial^2 X}{\partial t^2} = +c^2 \rho_0 \frac{\partial^2 X}{\partial x^2}.$$

Finally, cancel  $\rho_0$ :

$$\boxed{\frac{\partial^2 X}{\partial t^2} = c^2 \frac{\partial^2 X}{\partial x^2} .}$$

*Remark:* The thrust of this calculation has been to eliminate all variables but one. We chose to keep  $X$ , but could have chosen  $P_1$  instead, getting

$$\frac{\partial^2 P_1}{\partial t^2} = c^2 \frac{\partial^2 P_1}{\partial x^2} .$$

(Note that  $P_1$  is proportional to  $\partial X / \partial x$  by (II) and (I).) Also, the same equation is satisfied by the gas *velocity*,  $v(x, t) \equiv \partial X / \partial t$ .

## D'ALEMBERT'S SOLUTION

The wave equation,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},$$

can be solved by a special trick. (The rest of this course is devoted to other PDEs for which this trick does not work!)

Make a change of independent variables:

$$w \equiv x + ct, \quad z \equiv x - ct.$$

The dependent variable  $u$  is now regarded as a function of  $w$  and  $z$ . To be more precise one could write  $u(x, t) = \tilde{u}(w, z)$  (but I won't). We are dealing with a *different* function but the *same* physical quantity.

By the chain rule, acting upon any function we have

$$\frac{\partial}{\partial t} = \frac{\partial w}{\partial t} \frac{\partial}{\partial w} + \frac{\partial z}{\partial t} \frac{\partial}{\partial z} = c \frac{\partial}{\partial w} - c \frac{\partial}{\partial z},$$

$$\frac{\partial}{\partial x} = \frac{\partial w}{\partial x} \frac{\partial}{\partial w} + \frac{\partial z}{\partial x} \frac{\partial}{\partial z} = \frac{\partial}{\partial w} + \frac{\partial}{\partial z}.$$

Therefore,

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= c \left( \frac{\partial}{\partial w} - \frac{\partial}{\partial z} \right) \left[ c \left( \frac{\partial}{\partial w} - \frac{\partial}{\partial z} \right) u \right] \\ &= c^2 \left( \frac{\partial^2 u}{\partial w^2} - 2 \frac{\partial^2 u}{\partial w \partial z} + \frac{\partial^2 u}{\partial z^2} \right). \end{aligned}$$

Similarly,

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial w^2} + 2 \frac{\partial^2 u}{\partial w \partial z} + \frac{\partial^2 u}{\partial z^2}.$$

Thus the wave equation is

$$0 = \frac{1}{4} \left( \frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} \right) = \frac{\partial^2 u}{\partial w \partial z}.$$

This new equation is easily solved. We can write it in the form

$$\frac{\partial}{\partial w} \left( \frac{\partial u}{\partial z} \right) = 0.$$

Then it just says that  $\frac{\partial u}{\partial z}$  is a constant, as far as  $w$  is concerned. That is,

$$\frac{\partial u}{\partial z} = \gamma(z) \quad (\text{a function of } z \text{ only}).$$

Consequently,

$$u(w, z) = \int_{z_0}^z \gamma(\tilde{z}) d\tilde{z} + C(w),$$

where  $z_0$  is some arbitrary starting point for the indefinite integral. Note that the constant of integration will in general depend on  $w$ . Now since  $\gamma$  was arbitrary, its indefinite integral is an essentially arbitrary function too, and we can forget  $\gamma$  and just call the first term  $B(z)$ :

$$u(w, z) = B(z) + C(w).$$

(The form of the result is symmetrical in  $z$  and  $w$ , as it must be, since we could equally well have worked with the equation in the form  $\frac{\partial}{\partial z} \left( \frac{\partial u}{\partial w} \right) = 0$ .)

So, we have found the *general solution* of the wave equation to be

$$\boxed{u(x, t) = B(x - ct) + C(x + ct),}$$

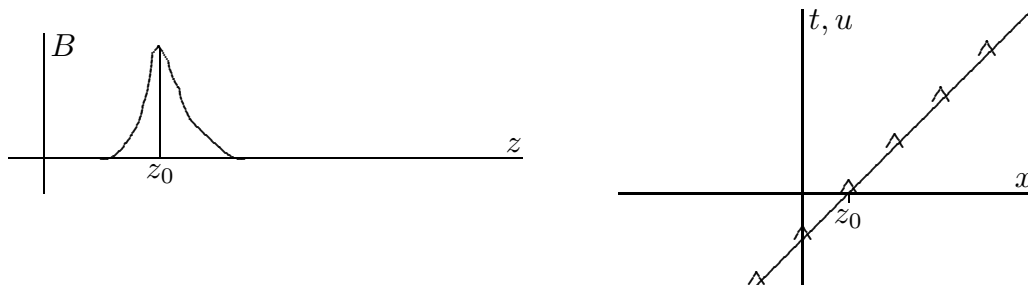
where  $B$  and  $C$  are arbitrary functions. (Technically speaking, we should require that the second derivatives  $B''$  and  $C''$  exist and are continuous, to make all our calculus to this point legal. However, it turns out that the d'Alembert formula remains meaningful and correct for choices of  $B$  and  $C$  that are much rougher than that.)

#### INTERPRETATION

What sort of function is  $B(x - ct)$ ? It is easiest to visualize if  $B(z)$  has a peak around some point  $z = z_0$ . Contemplate  $B(x - ct)$  as a function of  $x$  for a fixed  $t$ : It will have a peak in the neighborhood of a point  $x_0$  satisfying  $x_0 - ct = z_0$ , or

$$x_0 = z_0 + ct.$$

That is, the “bump” moves to the right with velocity  $c$ , keeping its shape exactly.



(Note that in the second drawing we have to plot  $u$  on the same axis as  $t$ . Such pictures should be thought of as something like a strip of movie film which we are forced to look at without the help of a projector.)\*

Similarly, the term  $C(x + ct)$  represents a wave pattern which moves rigidly to the *left* at the wave velocity  $-c$ . If both terms are present, and the functions are sharply peaked, we will see the two bumps collide and pass through each other. If the functions are not sharply peaked, the decomposition into left-moving and right-moving parts will not be so obvious to the eye.

### INITIAL CONDITIONS

In a concrete problem we are interested not in the most general solution of the PDE but in the particular solution that solves the problem! How much additional information must we specify to fix a unique solution? The *two arbitrary functions* in the general solution recalls the *two arbitrary constants* in the general solution of a second-order *ordinary* differential equation (ODE), such as

$$\frac{d^2u}{dt^2} + 4u = 0; \quad u(t) = B \sin(2t) + A \cos(2t).$$

In that case we know that the two constants can be related to two *initial conditions* (IC):

$$u(0) = A, \quad \frac{du}{dt}(0) = 2B.$$

Similarly, for the wave equation the two functions  $B(z)$  and  $C(w)$  can be related to initial data measured at, say,  $t = 0$ . (However, things will not be so simple for other second-order PDEs.)

Let's assume for the moment that our wave equation applies for *all* values of  $x$  and  $t$ :

$$-\infty < x < \infty, \quad -\infty < t < \infty.$$

We consider initial data at  $t = 0$ :

$$u(x, 0) = f(x), \quad \frac{\partial u}{\partial t}(x, 0) = g(x).$$

The d'Alembert solution implies

$$f(x) = B(x) + C(x), \quad g(x) = -cB'(x) + cC'(x).$$

---

\* In advanced physics, especially relativistic physics, it is standard to plot  $t$  on the *vertical* axis and  $x$  on the horizontal, even though for particle motion  $t$  is the independent variable and  $x$  the dependent one.

The second condition implies

$$-B(x) + C(x) = \int \frac{g(x)}{c} dx = G(x) + A,$$

where  $G$  is any antiderivative of  $g/c$ , and  $A$  is an unknown constant of integration. Solve these equations for  $B$  and  $C$ :

$$B(x) = \frac{1}{2}[f(x) - G(x) - A], \quad C(x) = \frac{1}{2}[f(x) + G(x) + A].$$

We note that  $A$  cancels out of the total solution,  $B(x - ct) + C(x + ct)$ . (Being constant, it qualifies as both left-moving and right-moving; so to this extent, the decomposition of the solution into left and right parts is ambiguous.) So we can set  $A = 0$  without losing any solutions. Now our expression for the solution in terms of the initial data is

$$u(x, t) = \frac{1}{2}[f(x + ct) + f(x - ct)] + \frac{1}{2}[G(x + ct) - G(x - ct)].$$

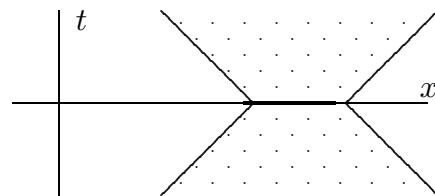
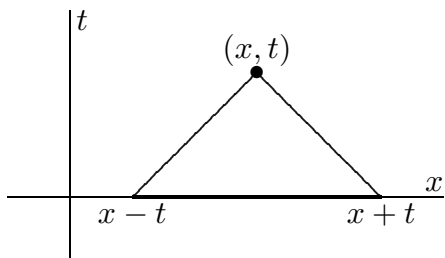
This is the first form of d'Alembert's fundamental formula. To get the second form, use the fundamental theorem of calculus to rewrite the  $G$  term as an integral over  $g$ :

$$u(x, t) = \frac{1}{2}[f(x + ct) + f(x - ct)] + \frac{1}{2c} \int_{x-ct}^{x+ct} g(w) dw.$$

This formula demonstrates that the value of  $u$  at a point  $(x, t)$  depends only on the part of the initial data representing “stuff” that has had time to reach  $x$  while traveling at speed  $c$  — that is, the data  $f(w, 0)$  and  $g(w, 0)$  on the *interval of dependence*

$$x - ct < w < x + ct \quad (\text{for } t > 0).$$

Conversely, any interval on the initial data “surface” (the line  $t = 0$ , in the two-dimensional case) has an expanding *region of influence* in space-time, beyond which its initial data are irrelevant. In other words, “signals” or “information” are carried by the waves with a finite maximum speed. These properties continue to hold for other wave equations (for example, in higher-dimensional space), even though in those cases the simple d'Alembert formula for the solution is lost and the waves no longer keep exactly the same shape as they travel.





## BOUNDARY CONDITIONS

In realistic problems one is usually concerned with only part of space (e.g, sound waves in a room). What happens to the waves at the edge of the region affects what happens inside. We need to specify this boundary behavior, in addition to initial data, to get a unique solution. To return to our physical example, if the sound waves are occurring in a closed pipe (of length  $L$ ), then the gas should be motionless at the ends:

$$X(0, t) = 0 = X(L, t).$$

Mathematically, these are called *Dirichlet boundary conditions* (BC). In contrast, if the pipe is open at one end, then to a good approximation the pressure at that point will be equal to the outside pressure,  $P_0$ . By our previous remark, this implies that the *derivative* of  $X$  vanishes at that end; for instance,

$$\frac{\partial X}{\partial x}(0, t) = 0$$

instead of one of the previous equations. This is called a *Neumann boundary condition*.

When a wave hits a boundary, it *reflects*, or “bounces off”. Let’s see this mathematically. Consider the interval  $0 < x < \infty$  and the Dirichlet condition

$$u(0, t) = 0.$$

Of course, we will have initial data,  $f$  and  $g$ , defined for  $x \in (0, \infty)$ .

We know that

$$u(x, t) = B(x - ct) + C(x + ct) \tag{1}$$

and

$$B(w) = \frac{1}{2}[f(w) - G(w)], \quad C(w) = \frac{1}{2}[f(w) + G(w)], \tag{2}$$

where  $f$  and  $cG' \equiv g$  are the initial data. However, if we try to calculate  $u$  from (1) for  $t > x/c$ , we find that (1) directs us to evaluate  $B(w)$  for *negative*  $w$ ; this is not defined in our present problem! To see what is happening, start at  $(x, t)$  and trace a right-moving ray backwards in time: It will run into the wall (the positive  $t$ -axis), not the initial-data surface (the positive  $x$ -axis).

Salvation is at hand through the boundary condition, which gives us the additional information

$$B(-ct) = -C(ct). \tag{3}$$

For  $t > 0$  this condition determines  $B$ (negative argument) in terms of  $C$ (positive argument). For  $t < 0$  it determines  $C$ (negative argument) in terms of  $B$ (positive argument). Thus  $B$  and  $C$  are uniquely determined for all arguments by (2) and (3) together.

In fact, there is a convenient way to represent the solution  $u(x, t)$  in terms of the initial data,  $f$  and  $g$ . Let us *define*  $f(x)$  and  $g(x)$  for negative  $x$  by requiring (2) to hold for negative values of  $w$  as well as positive. If we let  $y \equiv ct$ , (2) and (3) give (for all  $y$ )

$$f(-y) - G(-y) = -f(y) - G(y). \quad (4)$$

We would like to solve this for  $f(-y)$  and  $G(-y)$ , assuming  $y$  positive. But for that we need an independent equation (to get two equations in two unknowns). This is provided by (4) with negative  $y$ ; write  $y = -x$  and interchange the roles of right and left sides:

$$f(-x) + G(-x) = -f(x) + G(x). \quad (5)$$

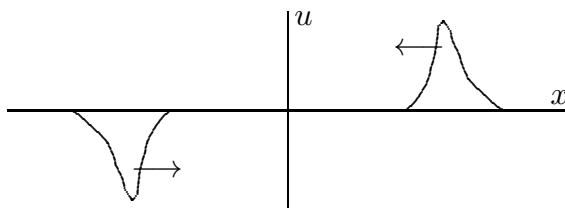
Rewrite (4) with  $y = +x$  and solve (4) and (5): For  $x > 0$ ,

$$f(-x) = -f(x), \quad G(-x) = G(x). \quad (6)$$

What we have done here is to define *extensions* of  $f$  and  $g$  from their original domain,  $x > 0$ , to the whole real line. The conditions (6) define the *odd extension* of  $f$  and the *even extension* of  $G$ . (It's easy to see that  $g = cG'$  is then odd, like  $f$ .) We can now solve the wave equation in all of  $\mathbf{R}^2$  ( $-\infty < x < \infty$ ,  $-\infty < t < \infty$ ) with these odd functions  $f$  and  $g$  as initial data. The solution is given by d'Alembert's formula,

$$u(x, t) = \frac{1}{2}[f(x + ct) + f(x - ct)] + \frac{1}{2}[G(x + ct) - G(x - ct)],$$

and it is easy to see that the boundary condition,  $u(0, t) = 0$ , is satisfied, because of the parity (evenness and oddness) of the data functions. Only the part of the solution in the region  $x > 0$  is physical; the other region is fictitious. In the latter region we have a “ghost” wave which is an *inverted mirror image* of the physical solution.



The calculation for Neumann conditions goes in very much the same way, leading to *even extensions* of  $f$  and  $g$ . The result is that the pulse reflects *without* turning upside down. Approximations to the “ideal” Dirichlet and Neumann boundary conditions are provided by a standard high-school physics experiment with Slinky<sup>TM</sup> springs. A small, light spring and a large, heavy one are attached end to end. When a wave traveling along the light spring hits the junction, the heavy spring remains almost motionless and the pulse reflects inverted. When the wave is in the heavy spring, the light spring serves merely to stabilize the apparatus; it carries off very little energy and barely constrains the motion of the end of the heavy spring. The pulse, therefore, reflects without inverting.

## TWO BOUNDARY CONDITIONS

Suppose that the spatial domain is  $0 < x < L$  with a Dirichlet condition at each end. The condition  $u(0, t) = 0$  can be treated by constructing odd and even extensions as before. The condition  $u(L, t) = 0$  implies, for all  $t$ ,

$$\begin{aligned} 0 &= B(L - ct) + C(L + ct) \\ &= \frac{1}{2}[f(L - ct) - G(L - ct)] + \frac{1}{2}[f(L + ct) + G(L + ct)]. \end{aligned} \quad (7)$$

Treating this equation as we did (4), we find an extension of  $f$  and  $G$  beyond the right end of the interval:

$$\begin{aligned} f(L + ct) &= -f(L - ct) = +f(-L + ct), \\ G(L + ct) &= G(L - ct) = G(-L + ct). \end{aligned}$$

(In more detail: Treat  $f(L+ct)$  and  $G(L+ct)$  with  $t > 0$  as the unknowns. Replacing  $t$  by  $-t$  in (7) gives two independent equations to be solved for them.) Finally, set  $ct = s + L$ :

$$f(s + 2L) = f(s), \quad G(s + 2L) = G(s) \quad (8)$$

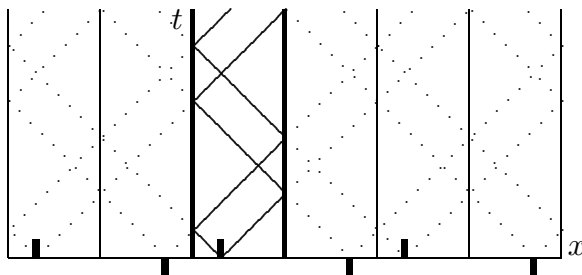
for all  $s$ . That is, the properly extended  $f$  and  $G$  (or  $g$ ) are *periodic* with period  $2L$ .

Here is another way to derive (8): Let's go back to the old problem with just one boundary, and suppose that it sits at  $x = L$  instead of  $x = 0$ . The basic geometrical conclusion can't depend on where we put the zero of the coordinate system: It must still be true that the extended data function is the odd (i.e., inverted) reflection of the original data through the boundary. That is, the value of the function at the point at a distance  $s$  to the left of  $L$  is minus its value at the point at distance  $s$  to the right of  $L$ . If the coordinate of the first point is  $x$ , then (in the case  $L > 0$ )  $s$  equals  $L - x$ , and therefore the coordinate of the second point is  $L + s = 2L - x$ . (This conclusion is worth remembering for future use: *The reflection of the point  $x$  through a boundary at  $L$  is located at  $2L - x$ .*) Therefore, the extended data function satisfies

$$f(x) = -f(2L - x).$$

In the problem with two boundaries, it also satisfies  $f(x) = -f(-x)$ , and thus  $f(2L - x) = f(x)$ , which is equivalent to the first half of (8) (and the second half can be proved in the same way).

The d'Alembert formula with these periodic initial data functions now gives a solution to the wave equation that satisfies the desired boundary and initial conditions. If the original initial data describe a single "bump", then the extended initial data describe an infinite sequence of image bumps, of alternating sign, as if space were filled with infinitely many parallel mirrors reflecting each other's images. Part of each bump travels off in each direction at speed  $c$ . What this really means is that the two wave pulses from the original, physical bump will suffer many reflections from the two boundaries. When a "ghost" bump penetrates into the physical region, it represents the result of one of these reflection events.



## HARSH FACTS OF LIFE

This PDE is not typical, even among *linear* ones.

1. For most linear PDEs, the waves (if indeed the solutions are wavelike at all) don't move without changing shape. They spread out. This includes higher-dimensional wave equations, and also the two-dimensional *Klein–Gordon equation*,

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - m^2 u,$$

which arises in relativistic quantum theory. (In homework, however, you are likely to encounter a partial extension of d'Alembert's solution to three dimensions.)

2. For most linear PDEs, it isn't possible to write down a simple general solution constructed from a few arbitrary functions.
3. For many linear PDEs, giving initial data on an open curve or surface like  $t = 0$  is not the most appropriate way to determine a solution uniquely. For example, *Laplace's equation*

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is the simplest of a class of PDEs called *elliptic* (whereas the wave equation is *hyperbolic*). For Laplace's equation the natural type of boundary is a closed curve, such as a circle, and only *one* data function can be required there.

## SEPARATION OF VARIABLES IN THE WAVE EQUATION

Let's again consider the wave equation on a finite interval with Dirichlet conditions (the vibrating string scenario):

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \tag{PDE}$$

where  $0 < x < L$  (but  $t$  is arbitrary),

$$u(0, t) = 0 = u(L, t), \quad (\text{BC})$$

$$u(x, 0) = f(x), \quad \frac{\partial u}{\partial t}(x, 0) = g(x). \quad (\text{IC})$$

During this first exposure to the method of variable separation, you should watch it as a “magic demonstration”. The reasons for each step and the overall strategy will be philosophized upon at length on future occasions.

We try the substitution

$$u(x, t) = X(x)T(t)$$

and see what happens. We have

$$\frac{\partial^2 u}{\partial t^2} = XT'', \quad \frac{\partial^2 u}{\partial x^2} = X''T,$$

and hence  $XT'' = c^2X''T$  from the PDE. Let's divide this equation by  $c^2XT$ :

$$\frac{T''}{c^2T} = \frac{X''}{X}.$$

This must hold for all  $t$ , and for all  $x$  in the interval. But the left side is a function of  $t$  only, and the right side is a function of  $x$  only. Therefore, the only way the equation can be true everywhere is that both sides are constant! We call the constant  $-K$ :

$$\frac{T''}{c^2T} = -K = \frac{X''}{X}.$$

Now the BC imply that

$$X(0)T(t) = 0 = X(L)T(t) \quad \text{for all } t.$$

So, either  $T(t)$  is identically zero, or

$$X(0) = 0 = X(L). \quad (*)$$

The former possibility would make the whole solution zero — an uninteresting, trivial case — so we ignore it. Therefore, we turn our attention to the ordinary differential equation satisfied by  $X$ ,

$$X'' + KX = 0, \quad (\dagger)$$

and solve it with the boundary conditions (\*).

Case 1:  $K = 0$ . Then  $X(x) = Ax + B$  for some constants. (\*) implies  $B = 0 = AL + B$ , hence  $A = 0 = B$ . This solution is also trivial.

Case 2:  $0 > K \equiv -\rho^2$ . Then

$$\begin{aligned} X(x) &= Ae^{\rho x} + Be^{-\rho x} \\ &= C \cosh(\rho x) + D \sinh(\rho x). \end{aligned}$$

The hyperbolic notation is the easier to work with in this situation. Setting  $x = 0$  in (\*), we see that  $C = 0$ . Then setting  $x = L$ , we get

$$0 = D \sinh(\rho L) \Rightarrow D = 0.$$

Once again we have run into the trivial solution. (The same thing happens if  $K$  is complex, but I won't show the details.)

Case 3:  $0 < K \equiv \lambda^2$ . This is our last hope. The solution is

$$X(x) = A \cos(\lambda x) + B \sin(\lambda x).$$

The boundary condition at  $x = 0$  gives  $A = X(0) = 0$ . The boundary condition at  $x = L$  gives

$$B \sin(\lambda L) = X(L) = 0.$$

We see that we can get a nontrivial solution if  $\lambda L$  is a place where the sine function equals zero. Well,  $\sin z = 0$  if and only if  $z = 0, \pi, 2\pi, \dots$ , or  $-\pi, -2\pi, \dots$ . That is,  $\lambda L = n\pi$  where  $n$  is an integer other than 0 (because we already excluded  $\lambda = 0$  as Case 1). Furthermore, we can assume  $n$  is positive, because the negative  $n$ s give the same functions as the positive ones, up to sign. Similarly, we can take  $B = 1$ , because multiplying a solution by a constant gives nothing new enough to be interesting. (For linear algebra students: We are interested only in solutions that are *linearly independent* of solutions we have already listed.)

In summary, we have found the solutions

$$X(x) = X_n(x) \equiv \sin \frac{n\pi x}{L}, \quad \sqrt{K} = \lambda_n \equiv \frac{n\pi}{L}, \quad n = 1, 2, \dots$$

The  $X$ s and  $\lambda$ s are called *eigenfunctions* and *eigenvalues* for the boundary value problem consisting of the ODE (†) and the BC (\*).

We still need to look at the equation for  $T$ :

$$T'' + c^2 \lambda^2 T = 0.$$

This, of course, has the general solution

$$T(t) = C \cos(c\lambda t) + D \sin(c\lambda t).$$

So, finally, we have found the *separated solution*

$$u_n(x, t) = \sin \frac{n\pi x}{L} \left( C \cos \frac{cn\pi t}{L} + D \sin \frac{cn\pi t}{L} \right)$$

for each positive integer  $n$ . (Actually, this is better thought of as *two independent separated solutions*, each with its arbitrary coefficient,  $C$  or  $D$ .)

## MATCHING INITIAL DATA

So far we have looked only at (PDE) and (BC). What initial conditions does  $u_n$  satisfy?

$$\begin{aligned} f(x) &= u(x, 0) = X(x)T(0) = C \sin(\lambda x), \\ g(x) &= \frac{\partial u}{\partial t}(x, 0) = X(x)T'(0) = c\lambda D \sin(\lambda x). \end{aligned}$$

Using trig identities, it is easy to check the consistency with D'Alembert's solution:

$$\begin{aligned} u(x, t) &= \sin(\lambda x)[C \cos(c\lambda t) + D \sin(c\lambda t)] \\ &= \frac{C}{2}[\sin \lambda(x - ct) + \sin \lambda(x + ct)] + \frac{D}{2}[\cos \lambda(x - ct) - \cos \lambda(x + ct)] \\ &= \frac{1}{2}[f(x + ct) + f(x - ct)] + \frac{1}{2}[G(x + ct) - G(x - ct)] \end{aligned}$$

where

$$G(z) = \frac{1}{c} \int^z g(x) dx = -D \cos(\lambda x) + \text{constant}.$$

The traveling nature of the  $x - ct$  and  $x + ct$  parts of the solution is barely noticeable, because they are spread out and superposed. The result is a standing vibration. It is called a *normal mode* of the system described by (PDE) and (BC).

But what if the initial wave profiles  $f(x)$  and  $g(x)$  aren't proportional to one of the eigenfunctions,  $\sin \frac{n\pi x}{L}$ ? The crucial observation is that both (PDE) and (BC) are *homogeneous linear* equations. That is,

- (1) the sum of two solutions is a solution;
- (2) a solution times a constant is a solution.

Therefore, any *linear combination* of the normal modes is a solution. Thus we know how to construct a solution with initial data

$$f(x) = \sum_{n=1}^N C_n \sin \frac{n\pi x}{L}, \quad g(x) = \sum_{n=1}^N \frac{cn\pi}{L} D_n \sin \frac{n\pi x}{L}.$$

This is still only a limited class of functions (all looking rather wiggly). But what about infinite sums?

$$f(x) = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi x}{L}, \quad \text{etc.}$$

**Fact:** Almost any function can be written as such a series of sines! That is what the next few weeks of the course is about. It will allow us to get a solution for any well-behaved  $f$  and  $g$  as initial data.

*Remark:* For discussion of these matters of principle, without loss of generality we can take  $L = \pi$ , so that

$$X_n(x) = \sin(nx), \quad \lambda_n = n.$$

We can always recover the general case by a change of variables,  $x = \pi\tilde{x}/L$ .

Before we leave the wave equation, let's take stock of how we solved it. I cannot emphasize too strongly that separation of variables always proceeds in *two steps*:

1. Hunt for separated solutions (normal modes). The assumption that the solution is separated ( $u_{\text{sep}} = X(x)T(t)$ ) is only for this intermediate calculation; most solutions of the PDE are *not* of that form. During this step we use only the *homogeneous* conditions of the problem — those that state that something is always equal to zero (in this case, (PDE) and (BC)).
2. Superpose the separated solutions (form a linear combination or an infinite series of them) and solve for the coefficients to match the data of the problem. In our example, “data” means the (IC). More generally, data equations are *nonhomogeneous linear* conditions: They have “nonzero right-hand sides”; adding solutions together yields a new solution corresponding to *different data*, the sum of the old data.

Trying to impose the initial conditions on an individual separated solution, rather than on a sum of them, leads to disaster! We will return again and again to the distinction between these two steps and the importance of not introducing a nonhomogeneous equation prematurely. Today is not the time for a clear and careful definition of “nonhomogeneous”, etc., but for some people a warning on this point in the context of this particular example may be more effective than the theoretical discussions to come later.



## Fourier Series

Now we need to take a theoretical excursion to build up the mathematics that makes separation of variables possible.

### PERIODIC FUNCTIONS

**Definition:** A function  $f$  is *periodic* with *period*  $p$  if

$$f(x + p) = f(x) \quad \text{for all } x.$$

*Examples and remarks:* (1)  $\sin(2x)$  is periodic with period  $2\pi$  — and also with period  $\pi$  or  $4\pi$ . (If  $p$  is a period for  $f$ , then an integer multiple of  $p$  is also a period. In this example the *fundamental* period — the smallest positive period — is  $\pi$ .) (2) The smallest common period of  $\{\sin(2x), \sin(3x), \sin(4x), \dots\}$  is  $2\pi$ . (Note that the fundamental periods of the first two functions in the list are  $\pi$  and  $2\pi/3$ , which are smaller than this common period.) (3) A constant function has every number as period.

The strategy of separation of variables raises this question:

Is every function with period  $2\pi$  of the form\*

$$(*) \quad \boxed{f(x) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]} ?$$

(Note that we could also write (\*) as

$$f(x) = \sum_{n=0}^{\infty} [a_n \cos(nx) + b_n \sin(nx)],$$

since  $\cos(0x) = 1$  and  $\sin(0x) = 0$ .)

More precisely, there are three questions:

1. What, exactly, does the infinite sum mean?
2. Given a periodic  $f$ , are there numbers  $a_n$  and  $b_n$  that make (\*) true?
3. If so, how do we calculate  $a_n$  and  $b_n$ ?

---

\* Where did the cosines come from? In the previous example we had only sines, because we were dealing with Dirichlet boundary conditions. Neumann conditions would lead to cosines, and periodic boundary conditions (for instance, heat conduction in a ring) would lead to both sines and cosines, as we'll see.

It is convenient to answer the last question first. That is, let's *assume* (\*) and then find formulas for  $a_n$  and  $b_n$  in terms of  $f$ . Here we make use of the ...

**Orthogonality relations:** If  $n$  and  $m$  are nonnegative integers, then

$$\begin{aligned} \int_{-\pi}^{\pi} \sin(nx) dx &= 0; \\ \int_{-\pi}^{\pi} \cos(nx) dx &= \begin{cases} 0 & \text{if } n \neq 0, \\ 2\pi & \text{if } n = 0; \end{cases} \\ \int_{-\pi}^{\pi} \sin(nx) \cos(mx) dx &= 0; \\ \int_{-\pi}^{\pi} \sin(nx) \sin(mx) dx &= \begin{cases} 0 & \text{if } n \neq m, \\ \pi & \text{if } n = m \neq 0; \end{cases} \\ \int_{-\pi}^{\pi} \cos(nx) \cos(mx) dx &= \begin{cases} 0 & \text{if } n \neq m, \\ \pi & \text{if } n = m \neq 0. \end{cases} \end{aligned}$$

*Proof:* These integrals are elementary, given such identities as

$$2 \sin \theta \sin \phi = \cos(\theta - \phi) - \cos(\theta + \phi).$$

Now multiply (\*) by  $\cos(mx)$  and integrate from  $-\pi$  to  $\pi$ . Assume temporarily that the integral of the series is the sum of the integrals of the terms. (To justify this we must answer questions 1 and 2.) If  $m \neq 0$  we get

$$\begin{aligned} \int_{-\pi}^{\pi} \cos(mx) f(x) dx &= a_0 \int_{-\pi}^{\pi} \cos(mx) dx \\ &+ \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos(mx) \cos(nx) dx + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi} \cos(mx) \sin(nx) dx \\ &= \pi a_m. \end{aligned}$$

We do similar calculations for  $m = 0$  and for  $\sin(mx)$ . The conclusion is: *If  $f$  has a Fourier series representation at all, then the coefficients must be*

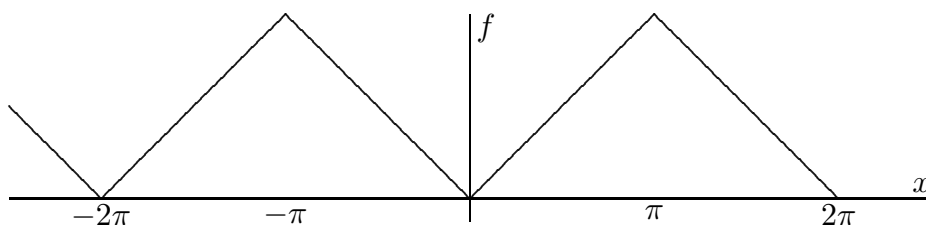
$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx, \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) f(x) dx, \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nx) f(x) dx. \end{aligned}$
--

Note that the first two equations can't be combined, because of an annoying factor of 2. (Some authors get rid of the factor of 2 by defining the coefficient  $a_0$  differently:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]. \quad (* \text{ NO } *)$$

In my opinion this is worse.)

*Example:* Find the Fourier coefficients of the function (“triangle wave”) which is periodic with period  $2\pi$  and is given for  $-\pi < x < \pi$  by  $f(x) = |x|$ .



$$\begin{aligned} \pi a_n &= \int_{-\pi}^{\pi} |x| \cos(nx) dx \\ &= \int_{-\pi}^0 (-x) \cos(nx) dx + \int_0^{\pi} x \cos(nx) dx. \end{aligned}$$

In the first term, let  $y = -x$ :

$$\begin{aligned} \pi a_n &= 2 \int_0^{\pi} x \cos(nx) dx \\ &= \frac{2}{n} \left[ x \sin(nx) \Big|_0^{\pi} - \int_0^{\pi} \sin(nx) dx \right] \\ &= 0 - \frac{2}{n} \frac{(-1)}{n} \cos(nx) \Big|_0^{\pi} \\ &= \frac{2}{n^2} (\cos(n\pi) - 1). \end{aligned}$$

Thus

$$a_n = \begin{cases} 0 & \text{if } n \text{ is even (and not 0),} \\ -\frac{4}{\pi n^2} & \text{if } n \text{ is odd.} \end{cases}$$

Similarly, one finds that  $a_0 = \frac{\pi}{2}$ . Finally,

$$\pi b_n = \int_{-\pi}^0 (-x) \sin(nx) dx + \int_0^{\pi} x \sin(nx) dx.$$

Here the first term equals  $\int_0^\pi y \sin(-ny) dy$ , but this is just the negative of the second term. So  $b_n = 0$ . (This will always happen when an *odd* integrand is integrated over an interval centered at 0.)

Putting the results together, we get

$$\begin{aligned} f(x) &\sim \frac{\pi}{2} + \sum_{k=0}^{\infty} \frac{-4}{\pi(2k+1)^2} \cos[(2k+1)x] \\ &= \frac{\pi}{2} - \frac{4}{\pi} \left[ \cos x + \frac{1}{9} \cos(3x) + \frac{1}{25} \cos(5x) + \dots \right]. \end{aligned}$$

(The symbol “ $\sim$ ” is a reminder that we have calculated the coefficients, but haven’t proved convergence yet. The important idea is that this “formal Fourier series” must have *something* to do with  $f$  even if it doesn’t converge, or converges to something other than  $f$ .)

It’s fun and informative to graph the first few *partial sums* of this series with suitable software, such as *Maple*. By taking enough terms of the series we really do get a good fit to the original function. Of course, with a *finite* number of terms we can never *completely* get rid of the wiggles in the graph, nor reproduce the sharp points of the true graph at  $x = n\pi$ .

#### FOURIER SERIES ON A FINITE INTERVAL

If  $f(x)$  is defined for  $-\pi < x \leq \pi$ , then it has a *periodic extension* to all  $x$ : just reproduce the graph in blocks of length  $2\pi$  all along the axis. That is,

$$f(x \pm 2\pi n) \equiv f(x) \quad \text{for any integer } n.$$

If  $f$  is continuous on  $-\pi < x \leq \pi$ , then the periodic extension is continuous if and only if

$$\lim_{x \downarrow -\pi} f(x) \equiv f(-\pi) = f(\pi) = \lim_{x \uparrow \pi} f(x).$$

(Here the operative equality (the target of “if and only if”) is the middle one. The left one is a definition, and the right one is a consequence of our continuity assumption. The notation  $\lim_{x \uparrow \pi}$  means the same as  $\lim_{x \rightarrow \pi^-}$ , etc.) This issue of continuity is important, because it influences how well the infinite Fourier series converges to  $f$ , as we’ll soon see.

The Fourier coefficients of the periodically extended  $f$ ,

$$\int_{-\pi}^{\pi} \cos(nx) f(x) dx \quad \text{and} \quad \int_{-\pi}^{\pi} \sin(nx) f(x) dx,$$

are completely determined by the values of  $f(x)$  in the original interval  $(-\pi, \pi]$  (or, for that matter, any other interval of length  $2\pi$  — all of which will give the same values for the integrals). Thus we think of a Fourier series as being associated with

- (1) an arbitrary function on a finite interval

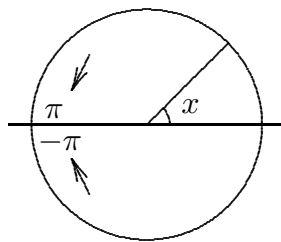
as well as

- (2) a periodic function on the whole real line.

Still another approach, perhaps the best of all, is to think of  $f$  as

- (3) an arbitrary function defined on a circle

with  $x$  as the *angle* that serves as coordinate on the circle. The angles  $x$  and  $x + 2\pi n$  represent the same point on the circle.



In particular,  $\pi$  and  $-\pi$  are the same point, no different in principle from any other point on the circle. Again,  $f$  (given for  $x \in (-\pi, \pi]$ ) qualifies as a *continuous function on the circle* only if  $f(-\pi) = f(\pi)$ . The behavior  $f(-\pi) \neq f(\pi)$  counts as a *jump discontinuity* in the theory of Fourier series.

*Caution:* The periodic extension of a function originally given on a finite interval is not usually the natural extension of the algebraic expression that defines the function on the original interval. The Fourier series belongs to the periodic extension, not the algebraic extension. For example, if  $f(x) = x^2$  on  $(-\pi, \pi]$ , its Fourier series is that of

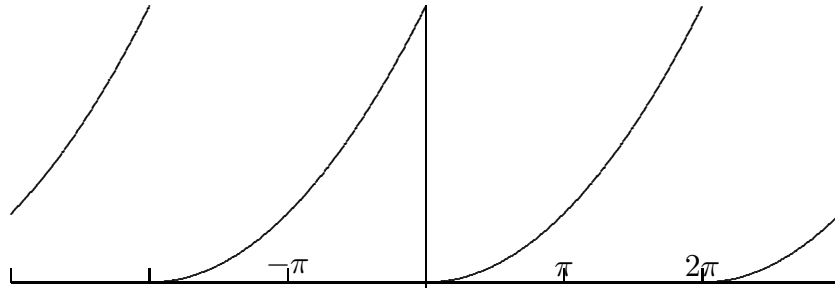


(axes not to scale!) and has nothing to do with the full parabola,

$$f(x) = x^2 \quad \text{for all } x.$$

The coefficients of this scalloped periodic function are given by integrals such as  $\int_{-\pi}^{\pi} \cos(mx) x^2 dx$ . If we were to calculate the integrals over some other interval

of length  $2\pi$ , say  $\int_0^{2\pi} \cos(mx) x^2 dx$ , then we would get the Fourier series of a very different function:



This does not contradict the earlier statement that the integration interval is irrelevant when you start with a function that is already periodic.

## EVEN AND ODD FUNCTIONS

An *even* function satisfies

$$f(-x) = f(x).$$

*Examples:*  $\cos$ ,  $\cosh$ ,  $x^{2n}$ .

An *odd* function satisfies

$$f(-x) = -f(x).$$

*Examples:*  $\sin$ ,  $\sinh$ ,  $x^{2n+1}$ .

In either case, the values  $f(x)$  for  $x < 0$  are determined by those for  $x > 0$  (or vice versa).

**Properties of even and odd functions** (schematically stated):

- (1) even + even = even; odd + odd = odd; even + odd = neither.

In fact, anything = even + odd:

$$f(x) = \frac{1}{2}[f(x) + f(-x)] + \frac{1}{2}[f(x) - f(-x)].$$

In the language of linear algebra, the even functions and the odd functions each form *subspaces*, and the vector space of all functions is their *direct sum*.

- (2) even  $\times$  even = even; odd  $\times$  odd = even; even  $\times$  odd = odd.

(3) (even)' = odd; (odd)' = even.

(4)  $\int$  odd = even;  $\int$  even = odd +  $C$ .

**Theorem:** If  $f$  is even, its Fourier series contains only cosines. If  $f$  is odd, its Fourier series contains only sines.

*Proof:* We saw this previously for an even example function. Let's work it out in general for the odd case:

$$\begin{aligned}\pi a_n &\equiv \int_{-\pi}^{\pi} f(x) \cos(nx) dx \\ &= \int_{-\pi}^0 f(x) \cos(nx) dx + \int_0^{\pi} f(x) \cos(nx) dx \\ &= \int_0^{\pi} f(-y) \cos(-ny) dy + \int_0^{\pi} f(x) \cos(nx) dx \\ &= 0.\end{aligned}$$

$$\begin{aligned}\pi b_n &\equiv \int_{-\pi}^{\pi} f(x) \sin(nx) dx \\ &= \int_{-\pi}^0 f(x) \sin(nx) dx + \int_0^{\pi} f(x) \sin(nx) dx \\ &= \int_0^{\pi} f(-y) \sin(-ny) dy + \int_0^{\pi} f(x) \sin(nx) dx \\ &= 2 \int_0^{\pi} f(x) \sin(nx) dx.\end{aligned}$$

This was for an odd  $f$  defined on  $(-\pi, \pi)$ . Given any  $f$  defined on  $(0, \pi)$ , we can extend it to an odd function on  $(-\pi, \pi)$ . Thus it has a Fourier series consisting entirely of sines:

$$\begin{aligned}f(x) &\sim \sum_{n=1}^{\infty} b_n \sin(nx) \\ \text{where } b_n &= \frac{2}{\pi} \int_0^{\pi} f(x) \sin(nx) dx \\ &\text{for odd } f \text{ on } -\pi < x < \pi \\ &\text{or any } f \text{ on } 0 < x < \pi.\end{aligned}$$

Similarly, the even extension gives a series of cosines for any  $f$  on  $0 < x < \pi$ . This series includes the constant term,  $n = 0$ , for which the coefficient formula has

an extra factor  $\frac{1}{2}$ . The formulas are

$$f(x) \sim \sum_{n=0}^{\infty} a_n \cos(nx)$$

where  $a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \cos(nx) dx$  for  $n > 0$ ,

$$a_0 = \frac{1}{\pi} \int_0^{\pi} f(x) dx$$

for even  $f$  on  $-\pi < x < \pi$   
or any  $f$  on  $0 < x < \pi$ .

For an interval of arbitrary length,  $L$ , we let  $x = \pi y/L$  and obtain

$$\bar{f}(y) \equiv f\left(\frac{\pi y}{L}\right) \sim \sum_{n=1}^{\infty} b_n \sin \frac{n\pi y}{L}$$

where  $b_n = \frac{2}{L} \int_0^L \bar{f}(y) \sin \frac{n\pi y}{L} dy$   
for odd  $\bar{f}$  on  $-L < y < L$   
or any  $\bar{f}$  on  $0 < y < L$ .

To keep the formulas simple, theoretical discussions of Fourier series are conducted for the case  $L = \pi$ ; the results for the general case then follow trivially.

**Summary:** Given an arbitrary function on an interval of length  $K$ , we can expand it in

- (1) sines *or* cosines of period  $2K$  (taking  $K = L$ , interval =  $(0, L)$ ),

*or*

- (2) sines *and* cosines of period  $K$  (taking  $K = 2L$ , interval =  $(-L, L)$ ).

In each case, the arguments of the trig functions in the series and the coefficient formulas are

$$\frac{m\pi x}{L}, \quad m = \text{integer.}$$

Which series to choose (equivalently, which extension of the original function) depends on the context of the problem; usually this means the *type of boundary conditions*.



## COMPLEX FOURIER SERIES

### A quick review of complex numbers:

$$i \equiv \sqrt{-1}.$$

Every complex number has the form  $z = x + iy$  with  $x$  and  $y$  real. To manipulate these, assume that  $i^2 = -1$  and all rules of ordinary algebra hold. Thus

$$(a + ib) + (c + id) = (a + c) + i(b + d);$$

$$(a + ib)(c + id) = (ac - bd) + i(bc + ad).$$

We write  $x \equiv \operatorname{Re} z$ ,  $y \equiv \operatorname{Im} z$ ;

$$|z| \equiv \sqrt{x^2 + y^2} = \text{modulus of } z;$$

$$z^* \equiv x - iy = \text{complex conjugate of } z.$$

Note that

$$(z_1 + z_2)^* = z_1^* + z_2^*, \quad (z_1 z_2)^* = z_1^* z_2^*.$$

Define

$$e^{i\theta} \equiv \cos \theta + i \sin \theta \quad (\theta \text{ real});$$

then

$$\begin{aligned} e^z &= e^{x+iy} \\ &= e^x e^{iy} \\ &= e^x (\cos y + i \sin y); \end{aligned}$$

$$|e^{i\theta}| = 1 \quad \text{if } \theta \text{ is real}; \quad e^{z+2\pi i} = e^z;$$

$$e^{i\pi} = -1, \quad e^{i\pi/2} = i, \quad e^{-i\pi/2} = e^{3\pi i/2} = -i = \frac{1}{i}, \quad e^{2\pi i} = e^0 = 1;$$

$$(e^{i\theta})^* = e^{-i\theta} = \frac{1}{e^{i\theta}}; \quad e^{-i\theta} = \cos \theta - i \sin \theta;$$

$$\boxed{\cos \theta = \frac{1}{2} (e^{i\theta} + e^{-i\theta}), \quad \sin \theta = \frac{1}{2i} (e^{i\theta} - e^{-i\theta}).}$$

*Remark:* Trig identities become trivial when expressed in terms of  $e^{i\theta}$ , hence easy to rederive. For example,

$$\begin{aligned} \cos^2 \theta &= \frac{1}{4} (e^{i\theta} + e^{-i\theta})^2 \\ &= \frac{1}{4} (e^{2i\theta} + 2 + e^{-2i\theta}) \\ &= \frac{1}{2} (\cos(2\theta) + 1). \end{aligned}$$

In the Fourier formulas (\*) for periodic functions on the interval  $(-\pi, \pi)$ , set

$$c_0 = a_0, \quad c_n = \frac{1}{2}(a_n - ib_n), \quad c_{-n} = \frac{1}{2}(a_n + ib_n).$$

The result is

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx},$$

where  $c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$

(Note that we are now letting  $n$  range through negative integers as well as nonnegative ones.) Notice that now there is *only one* coefficient formula. This is a major simplification!

Alternatively, the complex form of the Fourier series can be derived from *one* orthogonality relation,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} e^{-imx} dx = \begin{cases} 0 & \text{if } n \neq m, \\ 1 & \text{if } n = m. \end{cases}$$

As usual, we can scale these formulas to the interval  $(-L, L)$  by the variable change  $x = \pi y/L$ .

## CONVERGENCE THEOREMS

So far we've seen that we can solve the heat equation with homogenized Dirichlet boundary conditions and arbitrary initial data (on the interval  $[0, \pi]$ ), provided that we can express an arbitrary function  $g$  (on that interval) as an infinite linear combination of the eigenfunctions  $\sin(nx)$ :

$$g(x) = \sum_{n=1}^{\infty} b_n \sin nx.$$

Furthermore, we saw that if such a series exists, its coefficients must be given by the formula

$$b_n = \frac{2}{\pi} \int_0^{\pi} g(x) \sin nx dx.$$

So the burning question of the hour is: Does this Fourier sine series really converge to  $g(x)$ ?

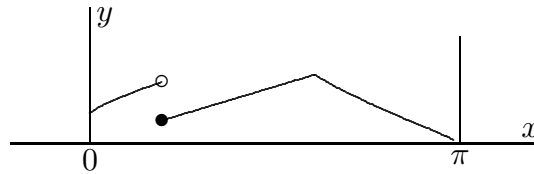
No mathematician can answer this question without first asking, "What kind of convergence are you talking about? And what technical conditions does  $g$  satisfy?" There are three standard convergence theorems, each of which states that certain technical conditions are sufficient to guarantee a certain kind of convergence. Generally speaking,

more smoothness in  $g$

$\iff$  more rapid decrease in  $b_n$  as  $n \rightarrow \infty$

$\iff$  better convergence of the series.

**Definition:**  $g$  is *piecewise smooth* if its derivative is piecewise continuous. That is,  $g'(x)$  is defined and continuous at all but a finite number of points (in the domain  $[0, \pi]$ , or whatever finite interval is relevant to the problem), and at those bad points  $g'$  has finite one-sided limits. (At such a point  $g$  itself is allowed to be discontinuous, but only the “finite jump” type of discontinuity is allowed.)



This class of functions is singled out, not only because one can rather easily prove convergence of their Fourier series (see next theorem), but also because they are a natural type of function to consider in engineering problems. (Think of electrical voltages under the control of a switch, or applied forces in a mechanical problem.)

**Pointwise Convergence Theorem:** If  $g$  is continuous and piecewise smooth, then its Fourier sine series converges at each  $x$  in  $(0, \pi)$  to  $g(x)$ . If  $g$  is piecewise smooth but not necessarily continuous, then the series converges to

$$\frac{1}{2}[g(x^-) + g(x^+)]$$

(which is just  $g(x)$  if  $g$  is continuous at  $x$ ). [Note that at the endpoints the series obviously converges to 0, regardless of the values of  $g(0)$  and  $g(\pi)$ . This zero is simply  $\frac{1}{2}[g(0^+) + g(0^-)]$  or  $\frac{1}{2}[g(\pi^+) + g(\pi^-)]$  for the odd extension!]

**Uniform Convergence Theorem:** If  $g$  is both continuous and piecewise smooth, and  $g(0) = g(\pi) = 0$ , then its Fourier sine series converges *uniformly* to  $g$  throughout the interval  $[0, \pi]$ .

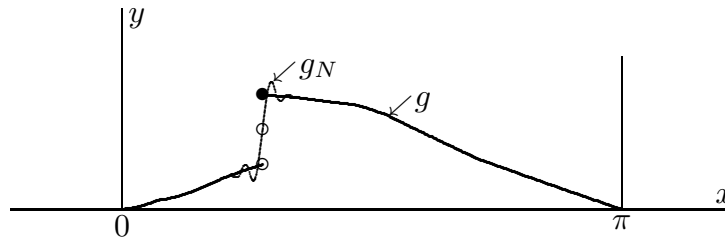
*Remarks:*

1. *Uniform convergence* means: For every  $\epsilon$  we can find an  $N$  so big that the partial sum

$$g_N(x) \equiv \sum_{n=1}^N b_n \sin(nx)$$

approximates  $g(x)$  to within an error  $\epsilon$  everywhere in  $[0, \pi]$ . The crucial point is that the same  $N$  works for *all*  $x$ ; in other words, you can draw a horizontal line,  $y = \epsilon$ , that lies completely above the graph of  $|g(x) - g_N(x)|$ .

2. In contrast, if the convergence is nonuniform (merely pointwise), then for each  $x$  we can take enough terms to get the error  $|g(x) - g_N(x)|$  smaller than  $\epsilon$ , but the  $N$  may depend on  $x$  as well as  $\epsilon$ . It is easy to see that if  $g$  is discontinuous, then uniform convergence is impossible, because the approximating functions  $g_N$  need a finite “time” to jump across the gap. There will always be points near the jump point where the approximation is bad.



It turns out that  $g_N$  develops “ears” or “overshoots” right next to the jump. This is called the *Gibbs phenomenon*.

3. For the same reason, the sine series can’t converge uniformly near an endpoint where  $g$  doesn’t vanish. An initial-value function which violated the condition  $g(0) = g(\pi) = 0$  would be rather strange from the point of view of the Dirichlet boundary value problem that gave rise to the sine series, since there we want  $u(0, t) = u(\pi, t) = 0$  and also  $u(x, 0) = g(x)$ !
4. If  $g$  is piecewise continuous, it can be proved that  $b_n \rightarrow 0$  as  $n \rightarrow \infty$ . (This is one form of the *Riemann–Lebesgue theorem*.) This is a key step in proving the pointwise convergence theorem.

If  $g$  satisfies the conditions of the uniform convergence theorem, then integration by parts shows that

$$b_n = \frac{2}{n\pi} \int_0^\pi g'(x) \cos(nx) dx,$$

and by another version of the Riemann–Lebesgue theorem this integral also approaches 0 when  $n$  is large, so that  $b_n$  falls off at  $\infty$  faster than  $n^{-1}$ . This additional falloff is “responsible” for the uniform convergence of the series. (This remark is as close as we’ll come in this course to proofs of the convergence theorems.)

5. There are continuous (but not piecewise smooth) functions whose Fourier series do *not* converge, but it is hard to construct an example! (See Appendix B.)

The third kind of convergence is related to . . .

**Parseval's Equation:** 
$$\int_0^\pi |g(x)|^2 dx = \frac{\pi}{2} \sum_{n=1}^{\infty} |b_n|^2.$$

(In particular, the integral converges if and only if the sum does.)

“*Proof*”: Taking convergence for granted, let's calculate the integral. (I'll assume that  $g(x)$  and  $b_n$  are real, although I've written the theorem so that it applies also when things are complex.)

$$\begin{aligned} \int_0^\pi |g(x)|^2 dx &= \int_0^\pi \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} b_n \sin(nx) b_m \sin(mx) dx \\ &= \int_0^\pi \sum_{n=1}^{\infty} b_n^2 \sin^2 nx dx \\ &= \frac{\pi}{2} \sum_{n=1}^{\infty} b_n^2. \end{aligned}$$

(The integrals have been evaluated by the orthogonality relations stated earlier. Only terms with  $m = n$  contribute, because of the orthogonality of the sine functions. The integral with  $m = n$  can be evaluated by a well known rule of thumb: The integral of  $\sin^2 \omega x$  over any integral number of quarter-cycles of the trig function is half of the integral of  $\sin^2 \omega x + \cos^2 \omega x$  — namely, the length of the interval, which is  $\pi$  in this case.)

There are similar Parseval equations for Fourier cosine series and for the full Fourier series on interval  $(-\pi, \pi)$ . In addition to its theoretical importance, which we can only hint at here, Parseval's equation can be used to evaluate certain numerical infinite sums, such as

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

(Work it out for  $g(x) = x$ .)

**Definition:**  $g$  is *square-integrable* on  $[0, \pi]$  if the integral in Parseval's equation converges:

$$\int_0^\pi |g(x)|^2 dx < \infty.$$

**$L^2$  (or Mean) Convergence Theorem:** If  $g$  is square-integrable, then the series converges *in the mean*:

$$\int_0^\pi |g(x) - g_N(x)|^2 dx \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

*Remarks:*

1. Recalling the formulas for the length and distance of vectors in 3-dimensional space,

$$|\vec{x}|^2 \equiv \sum_{n=1}^3 x_n^2, \quad |\vec{x} - \vec{y}|^2 \equiv \sum_{n=1}^3 (x_n - y_n)^2,$$

we can think of the Parseval integral as a measure of the “length” of  $g$ , and the integral in the theorem as a measure of the “distance” between  $g$  and  $g_N$ . (This geometrical way of thinking becomes very valuable when we consider general orthogonal basis functions later on.)

2. A function can be square-integrable without being piecewise smooth, or even bounded. Example:

$$g(x) \equiv \left(x - \frac{1}{2}\right)^{-\frac{1}{3}}.$$

Also (cf. Remark 5 above) a series can converge in the mean without converging pointwise (not to mention uniformly). This means that the equation

$$g(x) = \sum_{n=1}^{\infty} b_n \sin nx$$

must not be taken too literally in such a case — such as by writing a computer program to add up the terms for a fixed value of  $x$ . (The series will converge (pointwise) for “almost” all  $x$ , but there may be special values where it doesn’t.)

Prior to Fall 2000 this course spent about three weeks *proving* the convergence theorems and covering other aspects of the theory of Fourier series. (That material has been removed to make room for more information about PDEs, notably Green functions and the classification of PDEs as elliptic, hyperbolic, or parabolic.) Notes for those three weeks are attached as Appendix B.

## Fundamental Concepts: Linearity and Homogeneity

This is probably the most abstract section of the course, and also the most important, since the procedures followed in solving PDEs will be simply a bewildering welter of magic tricks to you unless you learn the general principles behind them. We have already seen the tricks in use in a few examples; it is time to extract and formulate the principles. (These ideas will already be familiar if you have had a good linear algebra course.)

### LINEAR EQUATIONS AND LINEAR OPERATORS

I think that you already know how to recognize linear and nonlinear equations, so let's look at some examples before I give the official definition of "linear" and discuss its usefulness.

#### Algebraic equations:

*Linear*

$$\begin{cases} x + 2y = 0, \\ x - 3y = 1 \end{cases}$$

*Nonlinear*

$$x^5 = 2x$$

#### Ordinary differential equations:

*Linear*

$$\frac{dy}{dt} + t^3 y = \cos 3t$$

*Nonlinear*

$$\frac{dy}{dt} = t^2 + e^y$$

#### Partial differential equations:

*Linear*

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

*Nonlinear*

$$\frac{\partial u}{\partial t} = \left( \frac{\partial u}{\partial x} \right)^2$$

What distinguishes the linear equations from the nonlinear ones? The most visible feature of the linear equations is that they involve the unknown quantity (the dependent variable, in the differential cases) *only to the first power*. The unknown does not appear inside transcendental functions (such as sin and ln), or

in a denominator, or squared, cubed, etc. This is how a linear equation is usually recognized by eye. Notice that there may be terms (like  $\cos 3t$  in one example) which don't involve the unknown at all. Also, as the same example term shows, there's no rule against nonlinear functions of the *independent* variable.

The formal definition of "linear" stresses not what a linear equation looks like, but the properties that make it easy to describe all its solutions. For concreteness let's assume that the unknown in our problem is a (real-valued) function of one or more (real) variables,  $u(x)$  or  $u(x, y)$ . The fundamental concept is not "linear equation" but "linear operator":

**Definition:** An operation,  $L$ , on functions is *linear* if it satisfies

$$L(u + v) = L(u) + L(v) \quad \text{and} \quad L(\lambda u) = \lambda L(u) \quad (*)$$

for all functions  $u$  and  $v$  and all numbers  $\lambda$ .

Examples of linear operations are

- differentiation of  $u$ :  $L(u) \equiv \frac{du}{dx}$ ,
- multiplication of  $u$  by a given function of  $x$ :  $L(u) \equiv x^2 u(x)$ ,
- evaluation of  $u$  at a particular value of  $x$ :  $L(u) \equiv u(2)$ ,
- integration of  $u$ :  $L(u) \equiv \int_0^1 u(x) dx$ .

In each example it's easy to check that  $(*)$  is satisfied, and we also see the characteristic *first-power* structure of the formulas (*without*  $u$ -independent terms this time). In each case  $L$  is a *function on functions*, a mapping which takes a function as input and gives as output either another function (as in the first two examples) or a number (as in the last two). Such a superfunction, considered as a mathematical object in its own right, is called an *operator*.

Now we can return to equations:

**Definition:** A *linear equation* is an equation of the form

$$L(u) = g,$$

where  $L$  is a linear operator,  $g$  is a "given" or "known" function (or number, as the case may be), and  $u$  is the unknown to be solved for.

So the possible  $u$ -independent terms enter the picture in the role of  $g$ . This leads to an absolutely crucial distinction:



## HOMOGENEOUS VS. NONHOMOGENEOUS EQUATIONS

**Definition:** A linear equation,  $L(u) = g$ , is *homogeneous* if  $g = 0$  (i.e., all terms in the equation are *exactly* of the first degree in  $u$ ); it is *nonhomogeneous* if  $g \neq 0$  (i.e., “constant” terms also appear).

In the second parenthetical clause, “constant” means independent of  $u$ . The “constant” term  $g$  may be a nontrivial function of the *independent* variable(s) of the problem.

Among our original examples, the linear ODE example was nonhomogeneous (because of the  $\cos 3t$ ) and the PDE example was homogeneous. The algebraic example is nonhomogeneous because of the 1. Here we are thinking of the system of simultaneous equations as a *single* linear equation in which the unknown quantity is a two-component vector,

$$\vec{u} \equiv \begin{pmatrix} x \\ y \end{pmatrix}.$$

The linear operator  $L$  maps  $\vec{u}$  onto another vector,

$$\vec{g} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

As you probably know, the system of equations can be rewritten in matrix notation as

$$\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The linear operator is described by the square matrix

$$M = \begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}.$$

In solving a differential equation we usually need to deal with initial or boundary conditions in addition to the equation itself. The main reason is that initial or boundary data need to be specified to give the problem a unique answer. Usually these conditions are themselves linear equations — for example, a standard initial condition for the heat equation:

$$u(0, x) = f(x).$$

Often the differential equation will be homogeneous but at least one of the boundary conditions will be nonhomogeneous. (The reverse situation also occurs.) Therefore, I think it’s helpful to introduce one more bit of jargon:

**Definitions:** A *linear problem* consists of one or more linear conditions (equations) to be satisfied by the unknown,  $u$ . A linear problem is *homogeneous* if all of its conditions are homogeneous, *nonhomogeneous* if one or more of the conditions are nonhomogeneous.

*Example A:* The ODE problem

$$u'' + 4u = 0, \quad u(0) = 1, \quad u'(0) = 0$$

is a nonhomogeneous linear problem. The ODE by itself is homogeneous, however.

*Example B:* The PDE problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + j(x), \quad u(x, 0) = 0, \quad u(0, t) = 0, \quad u(t, 1) = 0$$

is a nonhomogeneous linear problem. The boundary conditions *and* the initial condition are homogeneous, but the heat equation itself is nonhomogeneous in this case; the function  $j$  represents generation of heat inside the bar (perhaps by combustion or radioactivity), a possibility not considered in the discussion of the heat-conduction problem in Appendix A.

*Remark:* It is easy to see that every homogeneous linear equation has  $u = 0$  as a solution. (One proof:  $L(0) = L(u - u)$  (for any  $u$ ) =  $L(u) - L(u) = 0$ , QED.) Therefore, any homogeneous linear problem has 0 as a solution. Therefore, *if a linear problem has a unique solution and that solution is nontrivial (not just the 0 function), then that linear problem must be nonhomogeneous.* That is, an interesting, well-posed problem always has at least one nonhomogeneous condition.

## SOLVING LINEAR PROBLEMS

The importance of linear problems is that solving them is made easy by the *superposition principles* (which don't apply to nonlinear problems):

### Principles of Superposition:

1. A linear combination of solutions of a homogeneous problem is a new solution of that problem. That is, if  $L(u_1) = 0$  and  $L(u_2) = 0$ , then  $L(c_1 u_1 + c_2 u_2) = 0$  for any numbers  $c_1$  and  $c_2$  (and similarly for more than two solutions, and for more than one homogeneous linear equation defining the problem).

*Example:* Let **Problem 1** be the homogeneous ODE  $u'' + 4u = 0$ . Two solutions of this problem are

$$u_1 \equiv \cos 2x, \quad u_2 \equiv \sin 2x.$$

Then  $u = u_1 + 3u_2$ , for example, is also a solution. (In fact, we know that the *most general* solution is  $c_1u_1 + c_2u_2$  where the  $c$ 's are arbitrary constants. But for this we need a deeper existence-and-uniqueness theorem for second-order ODEs; it doesn't just follow from linearity.)

2. The sum of a solution of a nonhomogeneous problem and a solution of the *corresponding homogeneous problem* is a new solution of the original nonhomogeneous problem. ("Corresponding homogeneous problem" means the one with the same  $L$ 's, but with all  $g$ 's replaced by 0.)

*Example:* Let **Problem 2** be the nonhomogeneous equation  $u'' + 4u = e^x$ . One solution is  $u_p \equiv \frac{1}{5}e^x$ . (This has to be found by the method of undetermined coefficients, or by luck. Again, general principles of linearity by themselves can't solve the whole problem.) Now if we add a solution of Problem 1 we get a new solution of Problem 2:  $u_3 \equiv \frac{1}{5}e^x + \cos 2x$ .

3. The difference of two solutions of a nonhomogeneous problem is a solution of the corresponding homogeneous problem. Therefore, every solution of a nonhomogeneous problem can be obtained from *one particular* solution of that problem by adding some solution of the homogeneous problem.

*Example:* The general solution of Problem 2 is

$$u = \frac{1}{5}e^x + c_1 \cos 2x + c_2 \sin 2x.$$

4. The sum of solutions to two nonhomogeneous problems with the same  $L$ 's is a solution of a *new* nonhomogeneous problem, for which the  $g$ 's are the *sums* of the corresponding  $g$ 's of the two original problems. (Similarly for more than two nonhomogeneous problems.)

*Example 1:* The sum of two solutions of Problem 2,  $u_p$  and  $u_3$ , is  $z \equiv \frac{2}{5}e^x + \cos 2x$ , which is a solution of  $z'' + 4z = 2e^x$ . The important lesson to be learned from this example is that the right-hand side of this new equation is *not*  $e^x$ , the nonhomogeneous term of the two old equations. *Do not superpose solutions of a NONHOMOGENEOUS problem in the hope of getting a solution of that SAME problem.*

*Example 2:* Note that  $u_p$  is the unique solution of **Problem 3**:

$$u'' + 4u = e^x, \quad u(0) = \frac{1}{5}, \quad u'(0) = \frac{1}{5}.$$

Suppose that we really want to solve **Problem 4**:

$$u'' + 4u = e^x, \quad u(0) = 0, \quad u'(0) = 0.$$

Recalling Principles 2 and 3 as applied to the differential equation alone (not the initial conditions), we see that  $u = u_p + y$ , where  $y$  is some solution of  $y'' + 4y = 0$ . A moment's further thought shows that the correct  $y$  is the solution of **Problem 5**:

$$y'' + 4y = 0, \quad y(0) = -\frac{1}{5}, \quad y'(0) = -\frac{1}{5}.$$

A standard calculation shows that  $y = -\frac{1}{5} \cos 2x - \frac{1}{10} \sin 2x$ , and from this and  $u_p$  we can get the solution of Problem 4. (Of course, in solving such problems we usually don't write out Problem 5 as an intermediate step; the standard procedure is to impose the initial data of Problem 4 on the general solution found earlier. That is just a different way of organizing the same algebra. However, consciously splitting a nonhomogeneous problem into two nonhomogeneous problems, as I've demonstrated here for an ODE, is a common technique for solving PDEs.)

In summary, these principles provide the basic strategies for solving linear problems. If the problem is nonhomogeneous and complicated, you split it into simpler nonhomogeneous problems and add the solutions. If the solution is not unique, the nonuniqueness resides precisely in the possibility of adding a solution of the corresponding homogeneous problem. (In particular, if the original problem is homogeneous, then you seek the general solution as a linear combination of some list of basic solutions.) If the problem statement contains enough initial and boundary conditions, the solution will be unique; in that case, the only solution of the homogeneous problem is the zero function.

An important example application of this strategy is the solution of the heat-conduction problem in a bar with fixed end temperatures:\*

PDE: 
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},$$

IC: 
$$u(x, 0) = f(x),$$

BC: 
$$u(0, t) = T_1, \quad u(1, t) = T_2.$$

Here we have a homogeneous PDE, a nonhomogeneous initial condition, and two nonhomogeneous boundary conditions. The trick is to treat the two types of non-homogeneity separately. One writes  $u = v + w$ , where

---

\* See Appendix A, or Haberman's book.

- (1)  $v$  is to be a solution of the problem consisting of the PDE and the nonhomogeneous BC, with *no* particular IC assumed. It is possible to find a solution of this problem which is *independent of  $t$* :  $v(x, t) = V(x)$ .
- (2)  $w$  is to be a solution of the problem consisting of the PDE, the *homogeneous* Dirichlet boundary conditions

$$w(0, t) = 0, \quad w(1, t) = 0,$$

and the initial condition needed to make  $u$  satisfy the original IC. Namely,

$$w(x, 0) = f(x) - V(x).$$

It is very important that the only nonhomogeneity in this second problem is the IC. This makes it possible to solve for  $w$  by the method of separation of variables and then add the solutions without falling into the trap I warned you against earlier (Example 1). The solution is completed by finding the Fourier series of the function  $f - V$ .

The details of steps (1) and (2) are carried out in Appendix A.

For the processes of separating variables and calculating Fourier coefficients to work here, it was absolutely crucial to make the boundary conditions homogeneous first. In the calculation of normal modes, *no* nonhomogeneous conditions at all are imposed. The appropriate nonhomogeneous IC is imposed on a *superposition* ( $w$ ) of normal modes. Then still another term,  $v$ , is added to satisfy the nonhomogeneous BC.

One more time:

**Impose only HOMOGENEOUS conditions on normal modes (separated solutions).**

**Impose nonhomogeneous conditions only on a SUPERPOSITION (sum or integral) of normal modes.**

A related principle is

**Handle only one nonhomogeneity at a time!**

This principle is handled in practice by different strategies in different problems. Let's consider a doubly nonhomogeneous problem with the structure

$$L_1(u) = f_1, \quad L_2(u) = f_2.$$

The two principal strategies are these:

1. **“Zero out” the other condition.** Solve

$$\begin{aligned}L_1(u_1) &= f_1, & L_2(u_1) &= 0, \\L_1(u_2) &= 0, & L_2(u_2) &= f_2.\end{aligned}$$

Then  $u = u_1 + u_2$ .

Examples where this strategy is used include

- (a) treatment of the initial data  $u$  and  $\frac{\partial u}{\partial t}$  in the wave equation;
- (b) Laplace’s equation in a rectangle with boundary values given on two perpendicular sides.

2. **Temporarily ignore the other condition.** Solve  $L_1(u_1) = f_1$  and let  $L_2(u_1)$  be whatever it turns out to be, say  $L_2(u_1) \equiv h$ . Next solve

$$L_1(u_2) = 0, \quad L_2(u_2) = f_2 - h.$$

Then  $u = u_1 + u_2$ .

Examples where this strategy is used include

- (a) the method of undetermined coefficients for an ordinary differential equation with initial conditions;
- (b) finding a *steady-state solution* for the wave or heat equation with nonzero, but time-independent, boundary conditions.

## Moving into Higher Dimensions: The Rectangle

We will now work out a big example problem. It will break up into many small examples, which will demonstrate many of the principles we've talked about — often in a slightly new context.

### PROBLEM STATEMENT

We will consider heat conduction in a two-dimensional region, a rectangle. The ranges of the variables, therefore, will be

$$0 < x < a, \quad 0 < y < b, \quad t > 0.$$

Without loss of generality, we can assume that the variables have been scaled so that  $a = \pi$ .

The heat equation is

$$\text{PDE:} \quad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

Let us assume that the boundary conditions are

$$\text{BC}_1: \quad \frac{\partial u}{\partial x}(t, 0, y) = 0 = \frac{\partial u}{\partial x}(t, \pi, y),$$

$$\text{BC}_2: \quad u(t, x, 0) = p(x), \quad u(t, x, b) = q(x).$$

That is, the plate is insulated on the sides, and the temperature on the top and bottom edges is known and given by the functions  $p$  and  $q$ . Finally, there will be some initial temperature distribution

$$\text{IC:} \quad u(0, x, y) = f(x, y).$$

### STEADY-STATE SOLUTION

From our experience with the one-dimensional problem, we know that we must eliminate the nonhomogeneous boundary condition (BC<sub>2</sub>) before we can solve the initial-value problem by separation of variables! Fortunately,  $p$  and  $q$  are independent of  $t$ , so we can do this by the same technique used in one dimension: hunt for a time-independent solution of (PDE) and (BC),  $v(t, x, y) = V(x, y)$ , then consider the initial-value problem with homogeneous boundary conditions satisfied by  $u - v$ .

So, we first want to solve

$$\text{PDE:} \quad \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

$$\text{BC}_1: \quad \frac{\partial V}{\partial x}(0, y) = 0 = \frac{\partial V}{\partial x}(\pi, y),$$

$$\text{BC}_2: \quad V(x, 0) = p(x), \quad V(x, b) = q(x).$$

This is still a *partial* differential equation (namely, the two-dimensional *Laplace equation*). Furthermore, it still contains two nonhomogeneous conditions. Therefore, we split the problem again:

$$V = V_1 + V_2,$$

$$\begin{aligned} V_1(x, 0) &= p(x), & V_2(x, 0) &= 0, \\ V_1(x, b) &= 0, & V_2(x, b) &= q(x). \end{aligned}$$

Each  $V_j$  is supposed to satisfy Laplace's equation and  $(\text{BC}_1)$ .

*Remark:* This splitting is slightly different from the one involving the steady-state solution. In each subproblem here we have replaced every nonhomogeneous condition except one by its *corresponding homogeneous condition*. In contrast, for the steady-state solution we simply *discarded* the inconvenient nonhomogeneous condition, and later will modify the corresponding nonhomogeneous condition in the other subproblem to account for the failure of the steady-state solution to vanish on that boundary. Which of these techniques is best varies with the problem, but the basic principle is the same: Work with only one nonhomogeneous condition at a time, so that you can exploit the superposition principle correctly.

Let us solve for  $V_2$  by separation of variables:

$$V_{2\text{sep}}(x, y) = X(x)Y(y).$$

$$0 = X''Y + XY'' \Rightarrow -\frac{X''}{X} = \lambda = \frac{Y''}{Y}.$$

The boundary condition  $(\text{BC}_1)$  implies that

$$X'(0) = 0 = X'(\pi).$$

Therefore, up to a constant,

$$X(x) = \cos nx, \quad \lambda = n^2.$$



Now  $Y$  must be a solution of  $Y'' = n^2Y$  that vanishes at  $y = 0$ ; that is, up to a constant,

$$Y(y) = \sinh ny \quad \text{if } n \neq 0.$$

The case 0 must be treated separately:  $Y(y) = y$ . We have now taken care of three of the four boundaries. The remaining boundary condition is nonhomogeneous, and thus we cannot apply it to the individual separated solutions  $XY$ ; first we must add up the separated solutions with arbitrary coefficients:

$$V_2(x, y) = a_0y + \sum_{n=1}^{\infty} a_n \cos nx \sinh ny.$$

Now we must have

$$q(x) = a_0b + \sum_{n=1}^{\infty} a_n \cos nx \sinh nb.$$

This is a Fourier cosine series, so we solve for the coefficients by the usual formula:

$$a_n \sinh nb = \frac{2}{\pi} \int_0^{\pi} \cos nx q(x) dx \quad (n > 0).$$

Divide by  $\sinh nb$  to get a formula for  $a_n$ . For  $n = 0$  the Fourier formula lacks the factor 2, and we end up with

$$a_0 = \frac{1}{\pi b} \int_0^{\pi} q(x) dx.$$

This completes the solution for  $V_2$ .

Solving for  $V_1$  is exactly the same except that we need  $Y(b) = 0$  instead of  $Y(0) = 0$ . The appropriate solution of  $Y'' = n^2Y$  can be written as a linear combination of  $\sinh ny$  and  $\cosh ny$ , or of  $e^{ny}$  and  $e^{-ny}$ , but it is neater to write it as

$$Y(y) = \sinh(n(y - b)),$$

which manifestly satisfies the initial condition at  $b$  as well as the ODE. (Recall that hyperbolic functions satisfy trig-like identities, in this case

$$\begin{aligned} \sinh(n(y - b)) &= \cosh nb \sinh ny - \sinh nb \cosh ny \\ &= \frac{1}{2}e^{-nb} e^{ny} - \frac{1}{2}e^{nb} e^{-ny}, \end{aligned}$$

so the three forms are consistent.) Again the case  $n = 0$  is special:  $Y(y) = y - b$ . We now have

$$V_1(x, y) = A_0(y - b) + \sum_{n=1}^{\infty} A_n \cos nx \sinh n(y - b).$$

At  $y = 0$  this becomes

$$p(x) = -A_0 b - \sum_{n=1}^{\infty} A_n \cos nx \sinh nb.$$

Thus

$$A_n = -\frac{2}{\pi \sinh nb} \int_0^{\pi} \cos nx p(x) dx \quad (n > 0),$$

$$A_0 = -\frac{1}{\pi b} \int_0^{\pi} p(x) dx.$$

This completes the solution for  $V_1$  and hence for  $v(t, x, y)$ .

*Remark:* Since the boundary conditions at  $y = 0$  and  $y = b$  refer to the same variable, it was not really necessary to treat them separately. We could have separated variables in the problem [(Laplace PDE) + (BC<sub>1</sub>)] satisfied by the function  $V$ , getting

$$V_{\text{sep}}(x, y) = \cos nx Y(y), \quad Y'' = n^2 Y.$$

Then we could find the *general solution* of this last equation,

$$Y(y) = a_n \sinh ny + b_n \cosh ny$$

— or, better,

$$Y(y) = a_n \sinh ny + A_n \sinh n(y - b);$$

write the general superposition as a sum of these over  $n$ ; and then use the two nonhomogeneous boundary conditions (BC<sub>2</sub>) to determine the constants  $a_n$  and  $A_n$  in the summation.

This works because the nonhomogeneous conditions refer to parallel parts of the boundary. It definitely will not work for perpendicular edges! When in doubt, follow the injunction to deal with just one nonhomogeneity at a time.

#### HOMOGENEOUS PROBLEM

Next we're supposed to solve for  $w \equiv u - v$ , which must satisfy

PDE: 
$$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2},$$

BC<sub>1</sub>: 
$$\frac{\partial w}{\partial x}(t, 0, y) = 0 = \frac{\partial w}{\partial x}(t, \pi, y),$$

BC<sub>2</sub>: 
$$w(t, x, 0) = 0, \quad w(t, x, b) = 0,$$

IC: 
$$w(0, x, y) = f(x, y) - V(x, y) \equiv g(x, y).$$

Since there is only one nonhomogeneous condition, we can separate variables immediately:

$$\begin{aligned}w_{\text{sep}}(t, x, y) &= T(t)X(x)Y(y). \\T'XY &= TX''Y + TXY''. \\ \frac{T'}{T} &= \frac{X''}{X} + \frac{Y''}{Y} = -\lambda.\end{aligned}$$

(We know that  $\lambda$  is a constant, because the left side of the equation depends only on  $t$  and the right side does not depend on  $t$  at all. By analogy with the one-dimensional case we can predict that  $\lambda$  will be positive.) Since  $X''/X$  depends only on  $x$  and  $Y''/Y$  depends only on  $y$ , we can introduce *another* separation constant:

$$\frac{X''}{X} = -\mu, \quad \frac{Y''}{Y} = -\lambda + \mu.$$

The boundary conditions translate to

$$X'(0) = 0 = X'(\pi), \quad Y(0) = 0 = Y(b).$$

Thus for  $X$  we have the familiar solution

$$X(x) = \cos mx, \quad \mu = m^2.$$

Similarly, we must have

$$\begin{aligned}Y(y) &= \sin \frac{n\pi y}{b}, \quad -\lambda + \mu = -\frac{n^2\pi^2}{b^2} \\ \Rightarrow \lambda &= m^2 + \frac{n^2\pi^2}{b^2} \equiv \lambda_{mn}.\end{aligned}$$

Then

$$T(t) = e^{-\lambda t}.$$

(As usual in separation of variables, we have left out all the arbitrary constants multiplying these solutions. They will all be absorbed into the coefficients in the final Fourier series.)

We are now ready to superpose solutions and match the initial data. The most general solution of the homogeneous problem is a double infinite series,

$$w(t, x, y) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} c_{mn} \cos mx \sin \frac{n\pi y}{b} e^{-\lambda_{mn}t}.$$

The initial condition is

$$g(x, y) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} c_{mn} \cos mx \sin \frac{n\pi y}{b}. \quad (*)$$

To solve for  $c_{mn}$  we have to apply Fourier formulas twice:

$$\sum_{m=0}^{\infty} c_{mn} \cos mx = \frac{2}{b} \int_0^b \sin \frac{n\pi y}{b} g(x, y) dy;$$

$$c_{mn} = \frac{2}{\pi} \frac{2}{b} \int_0^{\pi} dx \int_0^b dy \cos mx \sin \frac{n\pi y}{b} g(x, y) \quad (m > 0),$$

$$c_{0n} = \frac{2}{\pi b} \int_0^{\pi} dx \int_0^b dy \sin \frac{n\pi y}{b} g(x, y).$$

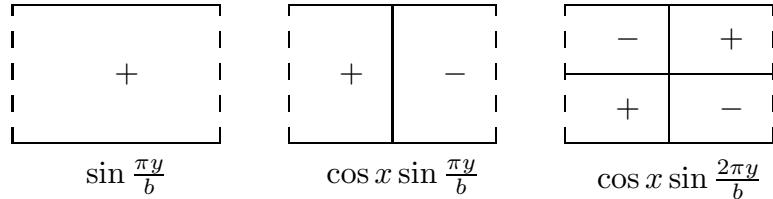
This completes the solution for  $w$ . Now we have the full solution to the original problem:

$$u(t, x, y) = w(t, x, y) + V(x, y).$$

Furthermore, along the way we have constructed a very interesting family of functions defined on the rectangle:

$$\phi_{mn}(x, y) \equiv \cos mx \sin \frac{n\pi y}{b}.$$

A few early members of the family look like this:



(Recall that  $\cos(0x) = 1$ .) The function is positive or negative in each region according to the sign shown. The function is zero on the solid lines and its normal derivative is zero along the dashed boundaries. The functions have these key properties for our purpose:

- They are *eigenvectors* of the Laplacian operator:

$$\left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \phi_{mn} = -\lambda_{mn} \phi_{mn}.$$

- *Completeness*: Any function (reasonably well-behaved) can be expanded as an infinite linear combination of them (the double Fourier series (\*)).
- *Orthogonality*: Each expansion coefficient  $c_{mn}$  can be calculated by a relatively simple integral formula, involving the corresponding eigenfunction  $\phi_{mn}$  only.

These functions form an *orthogonal basis* for the vector space of functions whose domain is the rectangle (more precisely, for the space  $L^2$  of square-integrable functions on the rectangle), precisely analogous to the orthogonal basis of eigenvectors for a symmetric matrix that students learn to construct in linear-algebra or ODE courses.

*Remark:* A complete treatment of convergence issues for the double Fourier series is not feasible here. We can say that if  $g(x, y)$  is very smooth, then the coefficients go to 0 fast as  $m$  or  $n \rightarrow \infty$ , and everything is OK. (More precisely, what needs to be smooth is the extension of  $g$  which is even and periodic in  $x$  and odd periodic in  $y$ . This places additional conditions on the behavior of  $g$  at the boundaries.) Also, if  $g$  is merely square-integrable, then the series converges in the mean, but not necessarily pointwise. (In that case the series for  $g$  can be used for certain theoretical purposes — e.g., inside the integrand of certain integrals — but an attempt to add it up on a computer is likely to lead to disaster.) However, when  $t > 0$  the series for  $w$  will converge nicely, even if  $g$  is rough, because the exponential factors make the terms decrease rapidly with  $m$  and  $n$ . This is a special feature of the heat equation: Because it describes a diffusive process, it drastically smooths out whatever initial data is fed into it.

#### THE CONSISTENCY CONDITION IN THE NEUMANN PROBLEM

Go back now to the steady-state problem and suppose that the boundary conditions on all four sides of the rectangle are of the normal-derivative type:

$$\text{PDE:} \quad \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

$$\text{BC}_1: \quad \frac{\partial V}{\partial x}(0, y) = f(y), \quad \frac{\partial V}{\partial x}(\pi, y) = g(y),$$

$$\text{BC}_2: \quad \frac{\partial V}{\partial y}(x, 0) = p(x), \quad \frac{\partial V}{\partial y}(x, b) = q(x).$$

Apply the two-dimensional version of Gauss's theorem:

$$\begin{aligned} 0 &= \int_0^\pi dx \int_0^b dy \nabla^2 V \\ &= \int_C \hat{\mathbf{n}} \cdot \nabla V \, ds \\ &= - \int_0^\pi f(y) \, dy + \int_0^\pi g(y) \, dy - \int_0^b p(x) \, dx + \int_0^b q(x) \, dx. \end{aligned}$$

Without even attempting to solve the problem, we can see that *there is no solution unless the net integral of the (outward) normal derivative data around the entire perimeter of the region is exactly equal to zero.*

This fact is easy to understand physically if we recall that this problem arose from a time-dependent problem of heat conduction, and that a Neumann boundary condition is a statement about *heat flow* out of the region concerned. If there is a net heat flow out of the region (and no heat source in the interior), then the rectangular object ought to be cooling off! It is not surprising that no *steady-state* solution can exist.

This existence problem is accompanied by a phenomenon of nonuniqueness, as often happens with linear equations. (Remember what happens to  $N$  equations in  $N$  unknowns when the determinant of the coefficient matrix is 0.) Suppose that the net heat flux *is* zero, and that we have found a solution,  $V$ , of the steady-state problem. Add a constant:  $V_*(x) \equiv V(x) + C$ . Since the constant function has zero Laplacian *and* zero normal derivatives all around,  $V_*$  is also a solution, no matter what  $C$  is. In the context of the original time-dependent heat problem, this ambiguity in the definition of the steady-state solution is merely a harmless nuisance: Just subtract  $C$  from the initial data ( $g(x, y)$ ) of the complementary problem with homogeneous boundary data, and the final solution will come out the same (unique).

## TWO TRICKS FOR RECTANGLES

The common lesson of these two examples is, “Just because you *can* expand an unknown solution in a Fourier series doesn’t mean that you *should*.” Sometimes a simply polynomial will do a better job.

### Retaining consistency in the Neumann problem

Consider Laplace’s equation in a rectangle with Neumann boundary conditions as above, and assume that the normal derivatives integrate to 0, so a solution should exist. Let’s reform the notation to make it more systematic:

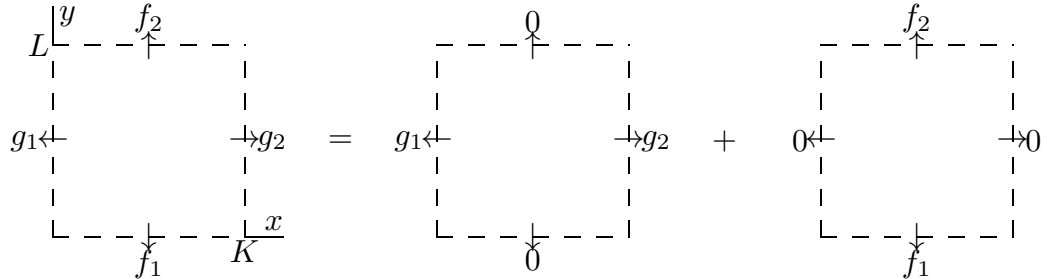
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

$$-\frac{\partial u}{\partial y}(x, 0) = f_1(x), \quad \frac{\partial u}{\partial y}(x, L) = f_2(x),$$

$$-\frac{\partial u}{\partial x}(0, y) = g_1(y), \quad \frac{\partial u}{\partial x}(K, y) = g_2(y),$$

with

$$\int_0^K [f_1(x) + f_2(x)] dx + \int_0^L [g_1(y) + g_2(y)] dy = 0.$$



Following the usual strategy, let's break up the problem into two, so that we have nonhomogeneous data in only one variable at a time. (The diagram indicates the resulting boundary equations.) But we have outfoxed ourselves.\* There is no reason why  $\int_0^K [f_1(x) + f_2(x)] dx$  and  $\int_0^L [g_1(y) + g_2(y)] dy$  should equal 0 individually, so in general the two subproblems will not have solutions. What to do?

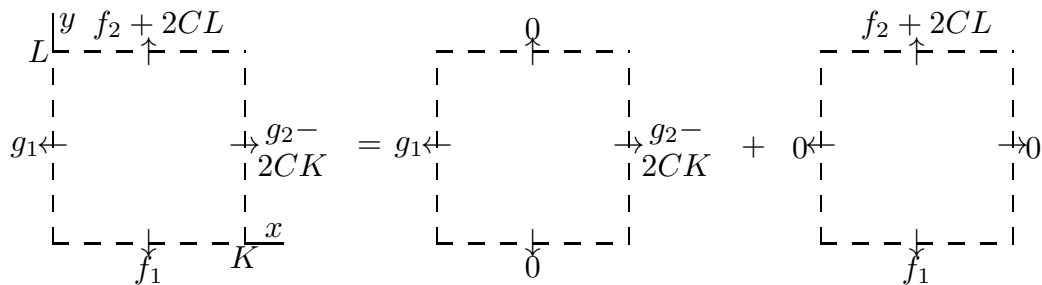
Here is a "magic rescue". The function  $V(x, y) \equiv x^2 - y^2$  satisfies  $\nabla^2 V = 0$  and

$$\frac{\partial V}{\partial x} = 2x = \begin{cases} 0 & \text{when } x = 0, \\ 2K & \text{when } x = K, \end{cases} \quad \frac{\partial V}{\partial y} = -2y = \begin{cases} 0 & \text{when } y = 0, \\ -2L & \text{when } y = L. \end{cases}$$

Let

$$C = -\frac{1}{2KL} \int_0^K [f_1(x) + f_2(x)] dx = +\frac{1}{2KL} \int_0^L [g_1(y) + g_2(y)] dy.$$

We would like to have a solution,  $u(x, y)$ , of the original problem with data  $f_1, f_2, g_1, g_2$ . Suppose for a moment that such a solution exists, and consider  $w \equiv u - CV$ . We see that  $\nabla^2 w = 0$  and that  $w$  satisfies Neumann boundary conditions shown in the next diagram, along with the obvious decomposition:



\* Pointed out by Juan Carcuz-Jerez, a student in Fall 2000 class.

We calculate

$$\int_0^L [g_1(y) + g_2(y) - 2CK] dy = 2CKL - 2CKL = 0,$$

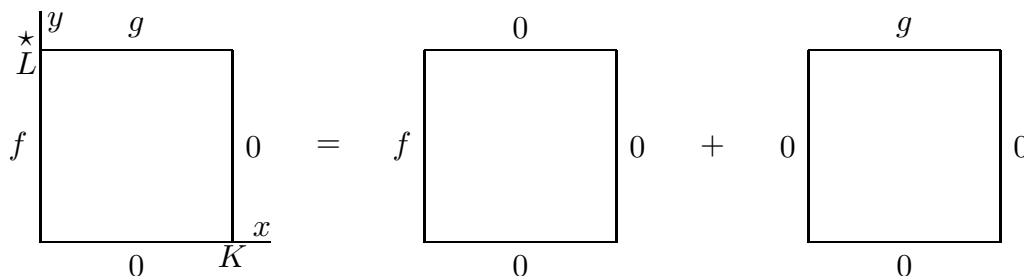
$$\int_0^K [f_1(x) + f_2(x) + 2CK] dx = -2CKL + 2CKL = 0.$$

Therefore, each of these subproblems *does* have a solution, which can be constructed as a Fourier cosine series in the usual way. (As usual in pure Neumann problems, the solutions are nonunique because an arbitrary constant could be added. Apart from that, the  $n = 0$  term in each cosine series is a function that is independent of the Fourier variable and linear in the other variable. (Try it and see!))

We can now define  $u = w + CV$  and observe that it solves the original Laplacian problem. (Hence it could serve as the steady-state solution for a related heat or wave problem.)

### Avoiding poor convergence at corners

Consider a Dirichlet problem for Laplace's equation with two nonhomogeneous conditions:



The two subproblems are solved by Fourier sine series in the usual way. Unless  $f(0) = 0 = f(L)$  and  $g(0) = 0 = g(K)$ , the solutions will demonstrate nonuniform convergence (and the Gibbs phenomenon). Suppose, however, that  $f$  and  $g$  are continuous (and piecewise smooth) and

$$f(0) = 0, \quad g(K) = 0, \quad f(L) = g(0) \neq 0.$$

Then the boundary data function is *continuous* all around the boundary, and one suspects that the optimal Fourier solution should be better behaved. The standard decomposition has introduced an *artificial discontinuity* at the corner marked “★” and thus a *spurious difficulty* of poor convergence.

The cure for this (admittedly relatively mild) disease is to consider

$$V(x, y) \equiv -g(0) \frac{y}{L} \frac{x - K}{K}.$$



We see that  $\nabla^2 V = 0$  and

$$V(\star) \equiv V(0, L) = g(0), \quad V(0, 0) = V(K, 0) = V(K, L) = 0.$$

Therefore,  $w \equiv u - V$  satisfies  $\nabla^2 w = 0$  with Dirichlet boundary data that *vanish at all four corners*. The problem for  $w$  can be decomposed into two subproblems in the usual way, and both of those will have uniformly convergent Fourier sine series.

More generally, any function of the form

$$V(x, y) = A + Bx + Cy + Dxy$$

is a solution of Laplace's equation. Given any continuous boundary data around a rectangle, the constants  $A, B, C, D$  can be chosen so that  $V$  matches the data exactly at all four corners. Then  $W \equiv u - V$  has continuous data that vanish at all four corners. By prudently subtracting off  $V$  before separating variables we get a better behaved Fourier solution. Of course, the double Fourier sine series for  $V(x, y)$  itself *would* exhibit nonuniform convergence, but there is no need here to decompose the simple polynomial function  $V$  in that way.)

## Fourier Transforms and Problems on Infinite Domains

Let's consider a heat conduction problem on a *semiinfinite* space interval:

$$\frac{1}{K} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad \text{for } 0 < x < \infty \text{ and } 0 < t < \infty.$$

The left end of the bar is insulated, so

$$\frac{\partial u}{\partial x}(0, t) = 0 \quad \text{for } 0 < t < \infty.$$

The initial temperature is

$$u(x, 0) = f(x) \quad \text{for } 0 < x < \infty.$$

When we try to solve this problem by separation of variables, we get as usual

$$X'' = -\lambda X, \quad T' = -\lambda K T.$$

If we set  $\lambda = \omega^2$ , the solutions of the  $X$  equation are

$$X(x) = A \cos(\omega x) + B \sin(\omega x).$$

The boundary condition  $X'(0) = 0$  forces  $B$  to be 0. We can choose  $A = 1$  and write the normal mode

$$u_\omega(x, t) = \cos(\omega x) e^{-\omega^2 K t}.$$

However, there is a major difference between this problem and the others we have considered: Since there is no second endpoint, there is no second boundary condition to determine the allowed values of  $\omega$ . Indeed, *all* nonnegative values of  $\omega$  are possible, and the complete solution  $u(x, t)$  satisfying the initial data will turn out to be an *integral* over these values, not a sum.\* That is why I have labeled the normal mode  $u_\omega$  instead of  $u_n$ ; *there is no integer variable  $n$  in this type of problem!*

Generally speaking, one has the correspondence

Finite interval  $\Rightarrow$  Fourier series (a sum);

Infinite interval  $\Rightarrow$  Fourier transform (an integral).

To see that this formulation is a slight oversimplification, note that a change of variable like  $y = \ln x$  can convert a finite interval into an infinite one  $[(0, 1)$  into  $(-\infty, 0)$ ]; obviously if a discrete sum is right in one case it will not become wrong in the other. (On the other hand, a Fourier series in  $x$  will no longer be a *Fourier* series in  $y$ , but something more general.) But this rule of thumb does apply to differential equations with *constant coefficients* and to some others. Note also that the interval referred to is one on which nonhomogeneous initial or boundary data are prescribed, not one where a homogeneous condition applies; we will see some examples of this distinction later.

---

\* You may wonder, then, why *complex* values of  $\omega$  are not also allowed. A completely satisfying answer is not possible at the level of technicality appropriate to this course, but a standard rule of thumb is that solutions that *increase* exponentially fast at infinity ( $\cosh(\kappa x)$  in this case) are not needed as eigenfunctions. We will soon see that the cosine functions by themselves are sufficient to represent all reasonable initial data.

INTUITIVE DERIVATION OF THE FOURIER TRANSFORM

It is easy to see how a Fourier series “becomes” an integral when the length of the interval goes to infinity. For this it is most convenient to use the complex-exponential form of the Fourier series. Recall that for a function on a finite interval of length  $2L$ , we have

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L},$$

$$c_n = \frac{1}{2L} \int_{-L}^L f(x) e^{-in\pi x/L} dx.$$

Let's write

$$k_n \equiv \frac{n\pi}{L}.$$

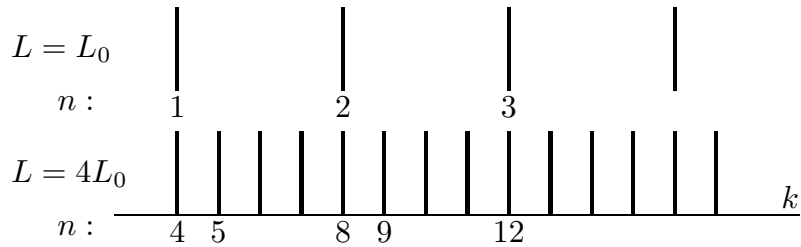
Then

$$f(x) = \sum_{k_n} c_n e^{ik_n x}.$$

The numbers  $k_n$  are called “frequencies” or “wave numbers”. As  $L$  increases, the frequencies become more closely spaced:

$$\Delta k_n \equiv \frac{(n+1)\pi}{L} - \frac{n\pi}{L} = \frac{\pi}{L}.$$

This suggests that for  $f$  defined on the whole real line,  $-\infty < x < \infty$ , *all* values of  $k$  should appear.



To make sense of the limit  $L \rightarrow \infty$ , we have to make a change of variable from  $n$  to  $k$ . Let

$$\hat{f}(k_n) \equiv L \sqrt{\frac{2}{\pi}} c_n.$$

Then

$$f(x) = \sqrt{\frac{\pi}{2}} \sum_{k_n} \frac{1}{L} \hat{f}(k_n) e^{ik_n x}$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{k_n} \hat{f}(k_n) e^{ik_n x} \Delta k_n,$$

$$\hat{f}(k_n) = \frac{1}{\sqrt{2\pi}} \int_{-L}^L f(x) e^{-ik_n x} dx.$$

As  $L \rightarrow \infty$  the first formula looks like a Riemann sum. In the limit we therefore expect

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk, \\ \hat{f}(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx. \end{aligned}$$

Note the surprising symmetry between these two formulas!  $\hat{f}$  is called the *Fourier transform* of  $f$ , and  $f$  is the *inverse Fourier transform* of  $\hat{f}$ .

### SINE AND COSINE TRANSFORMS

Of course, this does not solve our example problem. There the allowed functions were  $\cos(kx)$ , not  $e^{ikx}$ , and we were poised to expand an initial temperature distribution, defined for *positive  $x$  only*, in terms of them: If

$$f(x) \equiv u(x, 0) = \int_0^{\infty} A(k) \cos(kx) dk,$$

then

$$u(x, t) = \int_0^{\infty} A(k) \cos(kx) e^{-k^2 K t} dk$$

is the solution.

The way to get from exponentials to sines and cosines is basically the same as in finite Fourier series. First, note that the Fourier transformation we have derived (for  $-\infty < x < \infty$ ) can be rewritten in terms of  $\sin(kx)$  and  $\cos(kx)$  ( $0 \leq k < \infty$ ) in place of  $e^{ikx}$  ( $-\infty < k < \infty$ ). You can easily work out that the formulas are

$$f(x) = \int_0^{\infty} [A(k) \cos(kx) + B(k) \sin(kx)] dk,$$

$$A(k) = \frac{1}{\pi} \int_{-\infty}^{\infty} \cos(kx) f(x) dx,$$

$$B(k) = \frac{1}{\pi} \int_{-\infty}^{\infty} \sin(kx) f(x) dx.$$

This is seldom done in practical calculations with functions defined on  $(-\infty, \infty)$ , except by people with a strong hatred for complex numbers.

However, the trigonometric functions become very useful in calculations on a half-line (semiinfinite interval) with a boundary condition at the end. An arbitrary function on  $0 \leq x < \infty$  can be identified with its *even extension* to the whole real line. An *even* function has a Fourier transform consisting entirely of cosines (rather than sines), and the formula for the coefficient function can be written as an integral over just the positive half of the line:

$$f(x) = \int_0^{\infty} A(k) \cos(kx) dk,$$

$$A(k) = \frac{2}{\pi} \int_0^{\infty} \cos(kx) f(x) dx.$$

An equally common normalization convention splits the constant factor symmetrically between the two formulas:

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} A(k) \cos(kx) dk,$$

$$A(k) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(x) \cos(kx) dx.$$

Still other people put the entire factor  $\frac{2}{\pi}$  into the  $A \leftrightarrow F$  equation.\* In any case,  $A$  is called the *Fourier cosine transform* of  $f$ , and it's often given a notation such as  $\hat{f}_c(k)$  or  $F_C(k)$ .

It should now be clear how to finish the solution of the heat problem in the infinite bar with insulated left end.

Correspondingly, there is a *Fourier sine transform* related to *odd* extensions of functions. The formulas are the same except that  $\cos$  is replaced by  $\sin$  everywhere. The sine transform arises naturally in problems where the functions vanish at the boundary  $x = 0$ , and the cosine transform is appropriate when the derivative vanishes there (as we've seen).

## CONVERGENCE THEOREMS

Our derivation of the Fourier transformation formulas is not a proof that applying the two formulas in succession really will take you back to the function  $f$  from which you started; all the convergence theorems for Fourier series need to be reformulated and reproved for this new situation. In fact, since the integrals

---

\* Similar notational variations are found for the full (complex-exponential) Fourier transform.

are improper, the function  $f$  needs to satisfy some technical conditions before the integral  $\hat{f}$  will converge at all.

First, let's state the generalization to Fourier transforms of the pointwise convergence theorem for Fourier series. To get a true theorem, we have to make a seemingly fussy, but actually quite natural, technical condition on the function: Let's define a function with domain  $(-\infty, \infty)$  to be *piecewise smooth* if its restriction to every finite interval is piecewise smooth. (Thus  $f$  is allowed to have infinitely many jumps or corners, but they must not pile up in one region of the line.) The Fourier transform is defined by

$$\hat{f}(k) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx.$$

**Pointwise convergence theorem:** If  $f(x)$  is piecewise smooth, and

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty$$

( $f$  is *absolutely integrable*, or  $f \in L^1(-\infty, \infty)$ ), then:

- a)  $\hat{f}(k)$  is continuous.
- b)  $\hat{f}(k) \rightarrow 0$  as  $|k| \rightarrow \infty$  (but  $\hat{f}$  is not necessarily absolutely integrable itself). (This is a new version of the Riemann–Lebesgue theorem.)
- c) The inverse Fourier transform

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk$$

converges pointwise to  $\frac{1}{2}[f(x^+) + f(x^-)]$  (just like Fourier series).

The next theorem treats the variables  $x$  and  $k$  on a completely symmetrical basis.

**Mean convergence theorem:** If  $f(x)$  is sufficiently smooth to be integrated, and

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$$

( $f$  is *square-integrable*, or  $f \in L^2(-\infty, \infty)$ ), then:

- a)  $\hat{f}(k)$  is also square-integrable. (The integral defining  $\hat{f}(k)$  may not converge at every point  $k$ , but it will converge “in the mean”, just like the inversion integral discussed below.)

b) A Parseval equation holds:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}(k)|^2 dk.$$

(If you define  $\hat{f}$  so that the  $2\pi$  is kept all in one place, then this formula will not be so symmetrical.)

c) The inversion formula converges in the mean:

$$\lim_{\Lambda \rightarrow \infty} \int_{-\infty}^{\infty} dx |f(x) - f_{\Lambda}(x)|^2 = 0$$

where

$$f_{\Lambda}(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\Lambda}^{\Lambda} \hat{f}(k) e^{ikx} dk.$$

#### HOW FOURIER TRANSFORMS INTERACT WITH DERIVATIVES

**Theorem:** If the Fourier transform of  $f'$  is defined (for instance, if  $f'$  is in one of the spaces  $L^1$  or  $L^2$ , so that one of the convergence theorems stated above will apply), then the Fourier transform of  $f'$  is  $ik$  times that of  $f$ .

This can be seen either by differentiating

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk$$

with respect to  $x$ , or by integrating by parts in

$$\widehat{f'}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(x) e^{-ikx} dx$$

(at least if we assume that  $f(x) \rightarrow 0$  as  $x \rightarrow \pm\infty$ ).

Similarly, differentiation with respect to  $k$  corresponds to multiplication by  $-ix$ .

**Corollary:** If  $f'(x)$  exists and is in  $L^2$ , then  $k\hat{f}(k)$  is in  $L^2$ , and conversely, and similarly with  $x$  and  $k$  interchanged.

This is another instance of the principle that smoothness properties of  $f$  correspond to various degrees of rapid decrease of  $\hat{f}$  at  $\infty$ , and vice versa.

This differentiation property of Fourier transforms can be used to solve linear differential equations with constant coefficients. Consider the inhomogeneous ordinary differential equation

$$\frac{d^2 f}{dx^2} - \lambda^2 f(x) = g(x),$$

where  $\lambda^2 > 0$  and  $g$  is square-integrable and piecewise continuous. Take the Fourier transform of both sides:

$$(ik)^2 \hat{f}(k) - \lambda^2 \hat{f}(k) = \hat{g}(k).$$

Therefore,

$$\hat{f}(k) = \frac{-\hat{g}(k)}{k^2 + \lambda^2},$$

and hence

$$f(x) = \frac{-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\hat{g}(k)e^{ikx}}{k^2 + \lambda^2} dk.$$

(We won't know how to evaluate this integral, or even whether it can be done analytically, until we know what  $g$  is. Nevertheless, this is a definite formula for the solution. We'll return to this formula and press it a little farther later.)

What happened here is that the Fourier transformation converted a differential equation into an algebraic equation, which could be solved by elementary means. Our use of Fourier transforms (and Fourier series) to solve PDEs is really just an extension of this idea.

Once again, another way of looking at the calculation we have just done is as an analogue of *diagonalizing a matrix*. Suppose we want to solve the equation  $M\vec{x} = \vec{y}$ , where  $\vec{x}$  and  $\vec{y}$  are in  $\mathbf{R}^2$  and  $M$  is a  $2 \times 2$  matrix. If we can find a matrix  $U$  for which

$$M = U^{-1}DU, \quad D = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix},$$

then

$$M^{-1} = U^{-1}D^{-1}U, \quad D^{-1} = \begin{pmatrix} \frac{1}{m_1} & 0 \\ 0 & \frac{1}{m_2} \end{pmatrix}.$$

Then it is trivial to calculate  $\vec{y} = M^{-1}\vec{x}$ . In our ODE, the analogue of  $M$  is the differential operator  $\frac{d^2}{dx^2} - \lambda^2$  and the analogue of  $U$  is the Fourier transformation. We are using the fact that the functions  $e^{ikx}$  are *eigenvectors* of the differentiation operation  $\frac{d}{dx}$ , and hence of  $M$ .

You may (should!) object that the general solution of this ODE should contain two arbitrary constants. Indeed, the solution we have found is not the most general



one, but it is the only *square-integrable* one. (You can easily check that none of the solutions of the *associated homogeneous equation*,

$$\frac{d^2 f}{dx^2} - \lambda^2 f(x) = 0$$

(with  $\lambda^2 > 0$ ), are square-integrable, so adding one of them to our solution will give a solution of the inhomogeneous equation that is not in  $L^2$ .) The Fourier calculation in effect takes place entirely within the vector space  $L^2(-\infty, \infty)$  (although the eigenfunctions are not themselves members of that space).

### RELATION TO THE LAPLACE TRANSFORM

The foregoing may have reminded you of the Laplace-transform technique for solving ODEs. In fact, the two transforms are closely related.

Suppose  $f(x) = 0$  for  $x < 0$ . Then

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(x) e^{-ikx} dx.$$

Recall that the *Laplace transform* of  $f$  is

$$F(s) = \int_0^\infty f(x) e^{-sx} dx.$$

Allow  $s$  to be complex:

$$s = \sigma + ik, \quad \sigma \text{ and } k \text{ real.}$$

Then

$$\begin{aligned} F(s) &= \int_0^\infty f(x) e^{-\sigma x} e^{-ikx} dx \\ &= \sqrt{2\pi} \times \text{Fourier transform of } f(x)e^{-\sigma x} \quad (\sigma \text{ fixed}). \end{aligned}$$

For “most”  $f$ ’s,  $f(x)e^{-\sigma x}$  will be square-integrable if  $\sigma$  is sufficiently large, even if  $f$  itself is not square-integrable (e.g.,  $f = \text{polynomial for } x > 0$ ). To attain this result it was crucial that we cut  $f$  off below  $x = 0$ ; when we multiply by  $e^{-\sigma x}$ ,  $\sigma > 0$ , what we gain at  $x = +\infty$  we lose at  $-\infty$ . The Laplace transform (with *time* in the role of  $x$ ) is useful for solving initial-value problems, where the data and solution functions may not fall off to 0 as the time approaches  $+\infty$ , but negative values of time are not of interest. (In particular, the Laplace transform with respect

to time can be applied to nonhomogeneous boundary data that depend on time, so that the steady-state solution technique does not apply.)

The Fourier inversion formula for  $f e^{-\sigma x}$  says

$$f(x)e^{-\sigma x} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\sigma + ik) e^{ikx} dk,$$

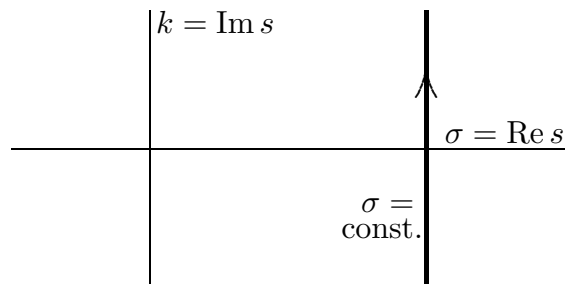
or

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\sigma + ik) e^{(\sigma+ik)x} dk.$$

In the exponent we recognize the complex variable  $s \equiv \sigma + ik$ . If we do a formal integration by substitution, taking  $ds = i dk$ , we get

$$f(x) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} F(s) e^{sx} ds.$$

In courses on complex analysis (such as Math. 407 and 601), it is shown that this integral makes sense as a line integral in the complex plane. It provides an *inversion formula for Laplace transforms*. In elementary differential-equations courses (such as Math. 308) no such formula was available; the only way to invert a Laplace transform was to “find it in the right-hand column of the table” — that is, to know beforehand that that function can be obtained as the direct Laplace transform of something else. The complex analysis courses also provide techniques for evaluating such integrals, so the number of problems that can be solved exactly by Laplace transforms is significantly extended.



In short, the Laplace transform is really the Fourier transform, extended to complex values of  $k$  and then rewritten in a notation that avoids complex numbers — until you want a formula to calculate the inverse transform, whereupon the complex numbers come back with a vengeance.

In this course we emphasize the use of the Fourier transform in solving partial differential equations. The Fourier transform also has important applications in signal processing and the analysis of data given as a function of a time variable. Here we take a quick look at some of the tools of that trade.

The Fourier transform of a product of functions is *not* the product of their Fourier transforms! Instead, it is easy to show that that transform is a certain integral involving the transforms of the two factors. This fact is most often used in the inverse direction, so that is how I'll state the formula:

**Convolution Theorem:** The inverse Fourier transform of  $\hat{f}_1(k)\hat{f}_2(k)$  is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_1(u)f_2(x-u) du \equiv f_1 * f_2.$$

This integral is called the *convolution* of  $f_1$  and  $f_2$ . Note that

$$f_1 * f_2 = f_2 * f_1,$$

although that is not immediately visible from the integral formula.

By manipulating the formulas defining the Fourier transform and its inverse, it is easy to show the following:

**Theorem:**

- (a) If  $g(x) \equiv f(-x)$ , then  $\hat{g}(k) = \hat{f}(-k)$ .
- (b) if  $g(x) \equiv f(-x)$  and  $f(x)$  is real-valued, then  $\hat{g}(k) = \hat{f}(k)^*$ .

Now take  $\hat{f}_1 = \hat{f}$  and  $\hat{f}_2 = \hat{f}^*$  in the convolution theorem and apply the theorem just stated:

**Corollary:** If  $f(x)$  is real-valued, then the Fourier transform of  $|\hat{f}(k)|^2$  is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u)f(u-x) du.$$

This integral is called the *autocorrelation function* of  $f$ , because it measures to what extent values of  $f$  at arguments displaced a distance  $x$  tend to coincide. The function  $|\hat{f}(k)|^2$  is called the *power spectrum* of  $f$ ; it measures the extent to which the signal in  $f$  is concentrated at frequency  $k$ .

As an application of the convolution theorem, return to the differential equation  $f'' - \lambda^2 f = g$  and the solution

$$f(x) = \frac{-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\hat{g}(k)e^{ikx}}{k^2 + \lambda^2} dk.$$

Suppose we knew that

$$-\frac{1}{k^2 + \lambda^2} = \hat{h}(k)$$

for some particular  $h(x)$ . Then we could write

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(x-t)g(t) dt$$

— thereby expressing the solution as a *single* integral instead of two (one to find  $\hat{g}$  and then one to find  $f$ ).

Can we find  $h$ ? Well, the most obvious way would be to evaluate the inverse Fourier transform,

$$h(x) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2 + \lambda^2} dk.$$

Unfortunately, one needs some theorems of complex analysis to evaluate this. Fortunately, I know the answer:

$$h(x) = -\frac{1}{\lambda} \sqrt{\frac{\pi}{2}} e^{-\lambda|x|} \quad (\text{if } \lambda > 0).$$

It can be verified by elementary means (see the next section) that this  $h$  satisfies

$$\hat{h}(k) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(x) e^{-ikx} dx = -\frac{1}{k^2 + \lambda^2}.$$

So we end up with

$$f(x) = -\frac{1}{2\lambda} \int_{-\infty}^{\infty} e^{-\lambda|x-t|} g(t) dt. \quad (*)$$

The function  $h$ , by the way, is called a *Green function* for this problem. It plays the same role as the matrix  $M^{-1}$  in the two-dimensional algebraic analogue.

Here is a way to check that (\*) is correct.

- (1) Find the general solution of  $f'' - \lambda^2 f = g$  by “variation of parameters” (see your differential equations textbook, or “Example 2” in the discussion of delta and Green functions below). The answer contains two arbitrary constants and some integrals that are beginning to look like (\*).
- (2) Determine the arbitrary constants by requiring that  $f$  be square-integrable. Then combine terms to get exactly (\*).

## THE UNCERTAINTY PRINCIPLE

If  $f(x)$  is sharply peaked, then  $\hat{f}(k)$  must be spread out; if  $\hat{f}(k)$  is sharply peaked, then  $f(x)$  must be spread out. A precise statement of this principle is:

$$\int_{-\infty}^{\infty} (x - x_0)^2 |f(x)|^2 dx \cdot \int_{-\infty}^{\infty} (k - k_0)^2 |\hat{f}(k)|^2 dk \geq \frac{1}{4} \|f\|^4$$

for any numbers  $x_0$  and  $k_0$ .

The proof appears in many textbooks of quantum mechanics, or in (for example) Dym and McKean, *Fourier Series and Integrals*, pp. 119–120. It uses the Schwarz inequality and the Parseval identity.

In quantum mechanics, when  $f$  is a *wave function* (in a unit system with Planck's constant  $\hbar = 1$ ),  $|f(x)|^2$  is the probability density for finding a particle at  $x$  and  $|\hat{f}(k)|^2$  is the probability density for measuring the particle's momentum to be  $k$ . The uncertainty principle is the mathematical reason why the position and momentum can't simultaneously be determined to arbitrary accuracy.

There is also a classical interpretation: Let  $x$  be time and  $f$  an electrical signal. Then  $\hat{f}(k)$  is its frequency spectrum. The uncertainty principle says that a pulse of brief duration must be composed of a wide spectrum of different frequencies; or that to qualify as truly monochromatic, a signal or wave must last for a long time.

## Green Functions

Here we will look at another example of how Fourier transforms are used in solving boundary-value problems. This time we'll carry the solution a step further, reducing the solution formula to a single integral instead of a double one.

### LAPLACE'S EQUATION IN THE UPPER HALF-PLANE

Let the ranges of the variables be

$$-\infty < x < \infty, \quad 0 < y < \infty.$$

Consider the equation

PDE: 
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

with the boundary data

BC: 
$$u(x, 0) = f(x).$$

This equation might arise as the steady-state problem for heat conduction in a *large* plate, where we know the temperature along one edge and want to simplify the problem by ignoring the effects of the other, distant edges. It could also arise in electrical or fluid-dynamical problems.

It turns out that to get a unique solution we must place one more condition on  $u$ : it must remain bounded as  $x$  or  $y$  or both go to infinity. (In fact, it will turn out that usually the solutions go to 0 at  $\infty$ .) Excluding solutions that grow at infinity seems to yield the solutions that are most relevant to real physical situations, where the region is actually finite. But it is the mathematics of the partial differential equation that tells us that to make the problem well-posed we do *not* need to *prescribe some arbitrary function* as the limit of  $u$  at infinity, as we needed to do in the case of finite boundaries.

Separating variables for this problem at first gives one a feeling of *déjà vu*:

$$u_{\text{sep}}(x, y) = X(x)Y(y) \Rightarrow 0 = X''Y + XY'';$$

$$-\frac{X''}{X} = \lambda = \frac{Y''}{Y};$$

write  $\lambda$  as  $k^2$ . The remaining steps, however, are significantly different from the case of the finite rectangle, which we treated earlier.

If  $\lambda \neq 0$ , the solution of the  $x$  equation can be

$$X(x) = e^{ikx},$$

where any  $k$  and its negative give the same  $\lambda$ . The condition of boundedness requires that  $k$  be real but does not further restrict it! Taking  $k = 0$  yields the only bounded solution with  $\lambda = 0$ . Therefore, we take the  $X$  in each separated solution to be  $e^{ikx}$  for some real  $k$ . The corresponding  $\lambda$  will be positive or zero.

Turning now to the  $y$  equation, we see that  $Y$  is some linear combination of  $e^{ky}$  and  $e^{-ky}$ . For boundedness we need the exponent to be negative, so we write

$$Y(y) = e^{-|k|y} (= e^{-\sqrt{\lambda}y})$$

to get an expression that's valid regardless of whether  $k$  is positive or negative.

We are now finished with the homogeneous conditions, so we're ready to superpose the separated solutions. Since  $k$  is a continuous variable, "superpose" in this case means "integrate", not "sum":

$$u(x, y) = \int_{-\infty}^{\infty} dk c(k) e^{ikx} e^{-|k|y}.$$

Here  $c(k)$  is an arbitrary function, which plays the same role as the arbitrary coefficients in previous variable separations. The initial condition is

$$f(x) = \int_{-\infty}^{\infty} dk c(k) e^{ikx}.$$

Comparing with the formula for the inverse Fourier transform, we see that  $c(k) = \frac{1}{\sqrt{2\pi}} \hat{f}(k)$ . That is,

$$c(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx.$$

In other words, the solution can be written

$$u(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \hat{f}(k) e^{ikx} e^{-|k|y}.$$

## A GREEN FUNCTION FOR LAPLACE'S EQUATION

We can get a simpler expression for  $u$  in terms of  $f$  by substituting the formula for  $\hat{f}$  into the one for  $u$ . But to avoid using the letter  $x$  to stand for two different

things in the same equation, we must first rewrite the definition of the Fourier transform using a different variable:

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dz e^{-ikz} f(z).$$

Then

$$u(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dz e^{ik(x-z)} e^{-|k|y} f(z).$$

We'll evaluate this multiple integral with the  $k$  integral on the inside. (This step requires some technical justification, but that is not part of our syllabus.) The inner integral is

$$\begin{aligned} \int_{-\infty}^{\infty} dk e^{ik(x-z)} e^{-|k|y} &= \int_{-\infty}^0 dk e^{ik(x-z)} e^{ky} + \int_0^{\infty} dk e^{ik(x-z)} e^{-ky} \\ &= \frac{e^{ik(x-z-iy)}}{i(x-z-iy)} \Big|_{-\infty}^0 + \frac{e^{ik(x-z+iy)}}{i(x-z+iy)} \Big|_0^{\infty} \\ &= \frac{1}{i(x-z-iy)} - \frac{1}{i(x-z+iy)} \\ &= \frac{2y}{(x-z)^2 + y^2}. \end{aligned}$$

Thus

$$u(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} dz \frac{y}{(x-z)^2 + y^2} f(z). \quad (*)$$

The function

$$G(x-z, y) \equiv \frac{1}{\pi} \frac{y}{(x-z)^2 + y^2}$$

is called a *Green function* for the boundary-value problem we started from. It is also called the *kernel* of the *integral operator*

$$u = G(f)$$

defined by (\*). The point of (\*) is that it gives the solution,  $u$ , as a function of the boundary data,  $f$ .

In principle, Green functions exist for the boundary-value problems on finite regions which we have solved earlier. However, in those cases the  $G$  is given by an infinite *sum* arising from the *Fourier series*, rather than the *integral* which expresses  $G$  in a Fourier-transform problem. Typically, such sums are harder to evaluate than the analogous integrals — which is why we have waited until now to introduce Green functions.



## GAUSSIAN INTEGRALS

The Green function for the *heat* equation on an infinite interval is derived from the Fourier-transform solution in much the same way. To do that we need a basic integral formula, which I'll now derive.

The integral in question is

$$H(x) \equiv \int_{-\infty}^{\infty} e^{ikx} e^{-k^2 t} dk,$$

where  $t$  is positive.

Note first that

$$\frac{d}{dk} e^{-k^2 t} = -2kt e^{-k^2 t}.$$

This will allow us to find a differential equation satisfied by  $H$ : From the definition we calculate

$$\begin{aligned} H'(x) &= \int_{-\infty}^{\infty} ik e^{ikx} e^{-k^2 t} dk \\ &= \frac{-i}{2t} \int_{-\infty}^{\infty} e^{ikx} \left( \frac{d}{dk} e^{-k^2 t} \right) dk \\ &= \frac{+i}{2t} \int_{-\infty}^{\infty} \left( \frac{d}{dk} e^{ikx} \right) e^{-k^2 t} dk \\ &= \frac{-x}{2t} \int_{-\infty}^{\infty} e^{ikx} e^{-k^2 t} dk \\ &= -\frac{x}{2t} H(x). \end{aligned}$$

Thus

$$\begin{aligned} \frac{H'}{H} &= -\frac{x}{2t}; \\ \ln H &= -\frac{x^2}{4t} + \text{const.}; \\ H &= C e^{-x^2/4t}. \end{aligned}$$

To find the constant we evaluate the integral for  $x = 0$ :

$$\begin{aligned} C &= H(0) \\ &= \int_{-\infty}^{\infty} e^{-k^2 t} dk \\ &= \frac{1}{\sqrt{t}} \int_{-\infty}^{\infty} e^{-q^2} dq, \end{aligned}$$

by the substitution  $q = k\sqrt{t}$ . But it is well known that

$$\int_{-\infty}^{\infty} e^{-q^2} dq = \sqrt{\pi},$$

because its square is

$$\begin{aligned} \iint_{\mathbf{R}^2} e^{-x^2} e^{-y^2} dx dy &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \\ &= 2\pi \int_0^{\infty} e^{-u} \frac{1}{2} du \\ &= \pi. \end{aligned}$$

So

$$C = \sqrt{\frac{\pi}{t}}.$$

Therefore, we have shown that  $H(x)$  is

$$\boxed{\int_{-\infty}^{\infty} e^{ikx} e^{-k^2 t} dk = \sqrt{\frac{\pi}{t}} e^{-x^2/4t}.$$

Now I leave it as an exercise\* to solve the initial-value problem for the heat equation for  $x \in (-\infty, \infty)$ :

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

$$u(0, x) = f(x), \tag{IC}$$

in analogy to our two previous Fourier-transform solutions. You should then find that the problem is solved by the Green function

$$G(t, x - z) \equiv \frac{1}{2\pi} H(x - z) = \frac{1}{\sqrt{4\pi t}} e^{-(x-z)^2/4t}.$$

Note also that the formula in the box is also useful for evaluating similar integrals with the roles of  $x$  and  $k$  interchanged. (Taking the complex conjugate of the formula, we note that the sign of the  $i$  in the exponent doesn't matter at all.)

---

\* Or peek at Haberman, Sec. 10.2.

## DELTA "FUNCTIONS"

The PDE problem defining any Green function is most simply expressed in terms of the *Dirac delta function*. This, written  $\delta(x - z)$  (also sometimes written  $\delta(x, z)$ ,  $\delta_z(x)$ , or  $\delta_0(x - z)$ ), is a make-believe function with these properties:

1.  $\delta(x - z) = 0$  for all  $x \neq z$ , and

$$\int_{-\infty}^{\infty} \delta(x - z) dx = 1.$$

2. **The key property:** For all continuous functions  $f$ ,

$$\int_{-\infty}^{\infty} \delta(x - z) f(x) dx = f(z).$$

Also,

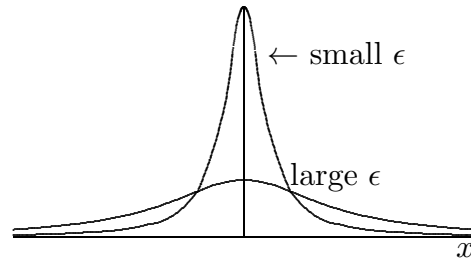
$$\int_a^b \delta(x - z) f(x) dx = \begin{cases} f(z) & \text{if } z \in (a, b), \\ 0 & \text{if } z \notin [a, b]. \end{cases}$$

3.  $\delta(x)$  is the limit of a family of increasingly peaked functions, each with integral 1:

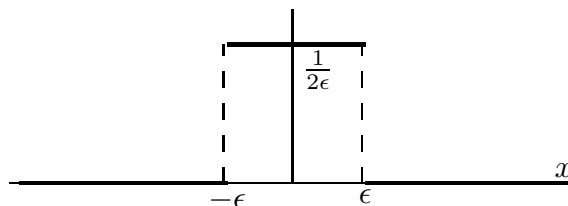
$$\delta(x) = \lim_{\epsilon \downarrow 0} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2}$$

$$\text{or } \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon\sqrt{\pi}} e^{-x^2/\epsilon^2}$$

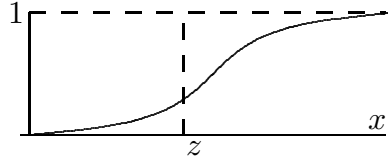
$$\text{or } \lim_{\epsilon \downarrow 0} d_\epsilon(x),$$



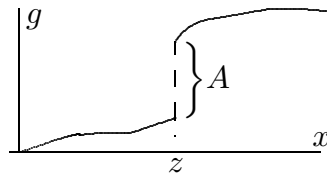
where  $d_\epsilon$  is a step function of the type drawn here:



4.  $\delta(x - z) = \frac{d}{dx}h(x - z)$ , where  $h(w)$  is the unit step function, or Heaviside function (equal to 1 for  $w > 0$  and to 0 for  $w < 0$ ). Note that  $h(t - z)$  is the limit as  $\epsilon \downarrow 0$  of a family of functions of this type:



GENERALIZATION OF 4: If  $g(x)$  has a jump discontinuity of size  $A$  at  $x = z$ , then its “derivative” contains a term  $A\delta(x - z)$ . ( $A$  may be negative.)

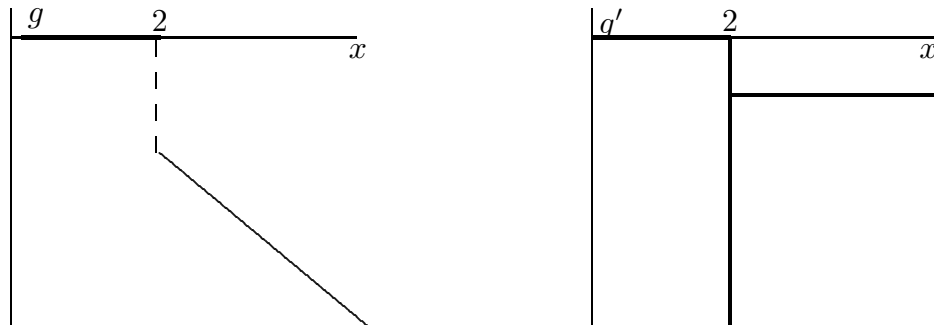


*Example:*

$$g(x) = \begin{cases} 0 & \text{for } x < 2, \\ -x & \text{for } x \geq 2 \end{cases} = -x h(x - 2).$$

Then

$$\begin{aligned} g'(x) &= -h(x - 2) - x h'(x - 2) \\ &= -h(x - 2) - 2\delta(x - 2). \end{aligned}$$



Consider

$$y'' + p(x)y' + q(x)y = A\delta(x - z).$$

We expect the solution of this equation to be the limit of the solution of an equation whose source term is a finite but very narrow and hard “kick” at  $x = z$ . The  $\delta$  equation is easier to solve than one with a finite peak.

The equation is taken to mean:

$$(1) \quad y'' + py' + qy = 0 \quad \text{for } x < z.$$

$$(2) \quad y'' + py' + qy = 0 \quad \text{for } x > z.$$

$$(3) \quad y \text{ is continuous at } z: \quad \lim_{x \downarrow z} y(x) = \lim_{x \uparrow z} y(x).$$

[Notational remarks:  $\lim_{x \downarrow z}$  means the same as  $\lim_{x \rightarrow z^+}$ ;  $\lim_{x \uparrow z}$  means  $\lim_{x \rightarrow z^-}$ . Also,  $\lim_{x \downarrow z} y(x)$  is sometimes written  $y(z^+)$ , and so on.]

$$(4) \quad \lim_{x \downarrow z} y'(x) = \lim_{x \uparrow z} y'(x) + A.$$

Conditions (3) and (4) tell us how to match solutions of (1) and (2) across the joint. Here is the reasoning behind them:

Assume (3) for the moment. Integrate the ODE from  $x = z - \epsilon$  to  $x = z + \epsilon$  (where  $\epsilon$  is very small):

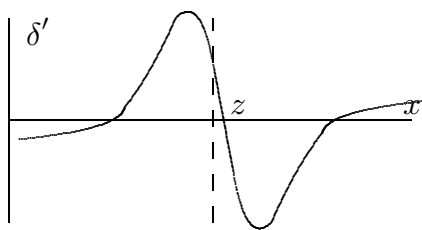
$$\int_{z-\epsilon}^{z+\epsilon} y'' dx + \int_{z-\epsilon}^{z+\epsilon} (py' + qy) dx = A \int_{z-\epsilon}^{z+\epsilon} \delta(x - z) dx$$

That is,

$$y'(z + \epsilon) - y'(z - \epsilon) + \text{small term } (\rightarrow 0 \text{ as } \epsilon \downarrow 0) = A.$$

In the limit  $\epsilon \rightarrow 0$ , (4) follows.

Now if  $y$  itself had a jump at  $z$ , then  $y'$  would contain  $\delta(x - z)$ , so  $y''$  would contain  $\delta'(x - z)$ , which is a singularity “worse” than  $\delta$ . (It is a limit of functions like the one in the graph shown here.) Therefore, (3) is necessary.



We can solve such an equation by finding the general solution on the interval to the left of  $z$  and the general solution to the right of  $z$ , and then matching the function and its derivative at  $z$  by rules (3) and (4) to determine the undetermined coefficients.

Consider the example

$$y'' = \delta(x - 1), \quad y(0) = 0, \quad y'(0) = 0.$$

For  $x < 1$ , we must solve the equation  $y'' = 0$ . The general solution is  $y = Ax + B$ , and the initial conditions imply then that

$$y = 0 \quad \text{for } x < 1.$$

For  $x > 1$ , we again must have  $y'' = 0$  and hence  $y = Cx + D$  (different constants this time). On this interval we have  $y' = C$ . To find  $C$  and  $D$  we have to apply rules (3) and (4):

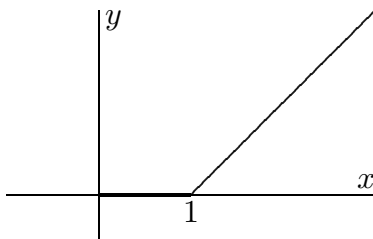
$$\begin{aligned} 0 &= y(1^-) = y(1^+) = C + D, \\ 0 + 1 &= y'(1^-) + 1 = y'(1^+) = C. \end{aligned}$$

That is,

$$\begin{aligned} C + D &= 0, \\ C &= 1. \end{aligned}$$

Therefore,  $C = 1$  and  $D = -1$ . Thus  $y(x) = x - 1$  for  $x > 1$ . The complete solution is therefore

$$y(x) = (x - 1) h(x - 1).$$



## DELTA FUNCTIONS AND GREEN FUNCTIONS

For lack of time, in this course we won't devote much attention to *nonhomogeneous* partial differential equations. (Haberman, however, discusses them extensively.) So far our nonhomogeneities have been initial or boundary data, not terms in the PDE itself. But problems like

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = \rho(t, x)$$

and

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = j(x, y),$$

where  $\rho$  and  $j$  are given functions, certainly do arise in practice. Often transform techniques or separation of variables can be used to reduce such PDEs to nonhomogeneous *ordinary* differential equations (a single ODE in situations of extreme symmetry, but more often an infinite family of ODEs).

Here I will show how the delta function and the concept of a Green function can be used to solve nonhomogeneous ODEs.

**Example 1:** *The Green function for the one-dimensional Dirichlet problem.* Let's start with an equation containing our favorite linear differential operator:

$$\frac{d^2 y}{dx^2} + \omega^2 y = f(x). \tag{*}$$

We require that

$$y(0) = 0, \quad y(\pi) = 0.$$

Here  $\omega$  is a positive constant, and  $f$  is a "known" but arbitrary function. Thus our solution will be a formula for  $y$  in terms of  $f$ . In fact, it will be given by a Green-function integral:

$$y(x) = \int_0^\pi G_\omega(x, z) f(z) dz,$$

where  $G$  is independent of  $f$  — but, of course, depends on the left-hand side of the differential equation (\*) and on the boundary conditions.

We can solve the problem for general  $f$  by studying the equation

$$\frac{d^2 y}{dx^2} + \omega^2 y = \delta(x - z) \tag{*_z}$$

(with the same boundary conditions). We will give the solution of (\*<sub>z</sub>) the name  $G_\omega(x, z)$ . Since

$$f(x) = \int_0^\pi \delta(x - z) f(z) dz$$

(for  $x$  in the interval  $(0, \pi)$ ) and since the operator on the left-hand side of  $(*)$  is linear, we expect that

$$y(x) \equiv \int_0^\pi G_\omega(x-z) f(z) dz$$

will be the solution to our problem! That is, since the operator is linear, it can be moved inside the integral (which is a limit of a sum) to act directly on the Green function:

$$\begin{aligned} \frac{d^2 y}{dx^2} + \omega^2 y &= \int_0^\pi \left( \frac{d^2}{dx^2} + \omega^2 \right) G_\omega(x-z) f(z) dz \\ &= \int_0^\pi \delta(x-z) f(z) dz \\ &= f(x), \end{aligned}$$

as desired. Furthermore, since  $G$  vanishes when  $x = 0$  or  $\pi$ , so does the integral defining  $y$ ; so  $y$  satisfies the right boundary conditions.

Therefore, the only task remaining is to solve  $(*_z)$ . We go about this with the usual understanding that

$$\delta(x-z) = 0 \quad \text{whenever } x \neq z.$$

Thus  $(*_z)$  implies

$$\frac{d^2 G_\omega(x, z)}{dx^2} + \omega^2 G_\omega(x, z) = 0 \quad \text{if } x \neq z.$$

Therefore, for some constants  $A$  and  $B$ ,

$$G_\omega(x, z) = A \cos \omega x + B \sin \omega x \quad \text{for } x < z,$$

and, for some constants  $C$  and  $D$ ,

$$G_\omega(x, z) = C \cos \omega x + D \sin \omega x \quad \text{for } x > z.$$

We do *not* necessarily have  $A = C$  and  $B = D$ , because the homogeneous equation for  $G$  is not satisfied when  $x = z$ ; that point separates the interval into two disjoint subintervals, and we have a *different* solution of the homogeneous equation on each. Note, finally, that the four unknown “constants” are actually functions of  $z$ : there is no reason to expect them to turn out the same for all  $z$ 's.

We need four equations to determine these four unknowns. Two of them are the boundary conditions:

$$0 = G_\omega(0, z) = A, \quad 0 = G_\omega(\pi, z) = C \cos \omega \pi + D \sin \omega \pi.$$



The third is that  $G$  is continuous at  $z$ :

$$A \cos \omega z + B \sin \omega z = G_\omega(z, z) = C \cos \omega z + D \sin \omega z.$$

The final condition is the one we get by integrating  $(*_z)$  over a small interval around  $z$ :

$$\frac{\partial}{\partial x} G_\omega(z^+, z) - \frac{\partial}{\partial x} G_\omega(z^-, z) = 1.$$

(Notice that although there is no variable “ $x$ ” left in this equation, the partial derivative with respect to  $x$  is still meaningful: it means to differentiate with respect to the first argument of  $G$  (before letting that argument become equal to the second one).) This last condition is

$$-\omega C \sin \omega z + \omega D \cos \omega z + \omega A \sin \omega z - \omega B \cos \omega z = 1.$$

One of the equations just says that  $A = 0$ . The others can be rewritten

$$\begin{aligned} C \cos \omega \pi + D \sin \omega \pi &= 0, \\ B \sin \omega z - C \cos \omega z - D \sin \omega z &= 0, \\ -\omega B \cos \omega z - \omega C \sin \omega z + \omega D \cos \omega z &= 1. \end{aligned}$$

This system can be solved by Cramer’s rule: After a grubby calculation, too long to type, I find that the determinant is

$$\begin{vmatrix} 0 & \cos \omega \pi & \sin \omega \pi \\ \sin \omega z & -\cos \omega z & -\sin \omega z \\ -\omega \cos \omega z & -\omega \sin \omega z & \omega \cos \omega z \end{vmatrix} = -\omega \sin \omega \pi.$$

If  $\omega$  is *not* an integer, this is nonzero, and so we can go on through additional grubby calculations to the answers,

$$B(z) = \frac{\sin \omega(z - \pi)}{\omega \sin \omega \pi},$$

$$C(z) = -\frac{\sin \omega z}{\omega},$$

$$D(z) = \frac{\cos \omega \pi \sin \omega z}{\omega \sin \omega \pi}.$$

Thus

$$G_\omega(x, z) = \frac{\sin \omega x \sin \omega(z - \pi)}{\omega \sin \omega \pi} \quad \text{for } x < z,$$

$$G_\omega(x, z) = \frac{\sin \omega z \sin \omega(x - \pi)}{\omega \sin \omega \pi} \quad \text{for } x > z.$$

(Reaching the last of these requires a bit more algebra and a trig identity.)

So we have found the Green function! Notice that it can be expressed in the unified form

$$G_\omega(x, z) = \frac{\sin \omega x_{<} \sin \omega(x_{>} - \pi)}{\omega \sin \omega \pi},$$

where

$$x_{<} \equiv \min(x, z), \quad x_{>} \equiv \max(x, z).$$

This symmetrical structure is very common in such problems.

Finally, if  $\omega$  is an integer, it is easy to see that the system of three equations in three unknowns has no solutions. It is no accident that these are precisely the values of  $\omega$  for which (\*)'s corresponding *homogeneous* equation,

$$\frac{d^2 y}{dx^2} + \omega^2 y = 0,$$

has solutions satisfying the boundary conditions. If the homogeneous problem has solutions (other than the zero function), then the solution of the nonhomogeneous problem (if it exists) must be nonunique, and we have no right to expect to find a formula for it! In fact, the *existence* of solutions to the nonhomogeneous problem also depends upon whether  $\omega$  is an integer (and also upon  $f$ ), but we don't have time to discuss the details here.

*Remark:* The algebra in this example could have been reduced by writing the solution for  $x > z$  as

$$G_\omega(x, z) = E \sin \omega(x - \pi).$$

(That is, we build the boundary condition at  $\pi$  into the formula by a clever choice of basis solutions.) Then we would have to solve merely two equations in two unknowns ( $B$  and  $E$ ) instead of a  $3 \times 3$  system.

**Example 2:** *Variation of parameters in terms of delta and Green functions.* Let's go back to the general second-order linear ODE,

$$y'' + p(x)y' + q(x)y = f(x),$$

and construct the solution satisfying

$$y(0) = 0, \quad y'(0) = 0.$$

As before, we will solve

$$\frac{\partial^2}{\partial x^2} G(x, z) + p(x) \frac{\partial}{\partial x} G(x, z) + q(x) G(x, z) = \delta(x - z)$$

with those initial conditions, and then expect to find  $y$  in the form

$$y(x) = \int G(x, z) f(z) dz.$$

It is not immediately obvious what the limits of integration should be, since there is no obvious “interval” in this problem.

Assume that two linearly independent solutions of the homogeneous equation

$$y'' + p(x)y' + q(x)y = 0$$

are known; call them  $y_1(x)$  and  $y_2(x)$ . Of course, until we are told what  $p$  and  $q$  are, we can't write down exact formulas for  $y_1$  and  $y_2$ ; nevertheless, we can solve the problem in the general case — getting an expression for  $G$  in terms of  $y_1$  and  $y_2$ , whatever they may be.

Since  $G$  satisfies the homogeneous equation for  $x \neq z$ , we have

$$G(x, z) = \begin{cases} A(z)y_1(x) + B(z)y_2(x) & \text{for } x < z, \\ C(z)y_1(x) + D(z)y_2(x) & \text{for } x > z. \end{cases}$$

As before we will get four equations in the four unknowns, two from initial data and two from the continuity of  $G$  and the prescribed jump in its derivative at  $z$ . Let us consider only the case  $z > 0$ . Then the initial conditions

$$G(0, z) = 0, \quad \frac{\partial}{\partial x}G(0, z) = 0$$

force  $A = 0 = B$ . The continuity condition, therefore, says that  $G(z, z) = 0$ , or

$$C(z)y_1(z) + D(z)y_2(z) = 0. \tag{1}$$

The jump condition

$$\frac{\partial}{\partial x}G(z^+, z) - \frac{\partial}{\partial x}G(z^-, z) = 1$$

now becomes

$$C(z)y_1'(z) + D(z)y_2'(z) = 1. \tag{2}$$

Solve (1) and (2): The determinant is the Wronskian

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1y_2' - y_2y_1' \equiv W(z).$$

Then

$$C = -\frac{y_2}{W}, \quad D = \frac{y_1}{W}.$$

Thus our conclusion is that (for  $z > 0$ )

$$G(x, z) = \begin{cases} 0 & \text{for } x < z, \\ \frac{1}{W(z)}(y_1(z)y_2(x) - y_2(z)y_1(x)) & \text{for } x > z. \end{cases}$$

Now recall that the solution of the original ODE,

$$y'' + p(x)y' + q(x)y = f(x),$$

was supposed to be

$$y(x) = \int G(x, z) f(z) dz.$$

Assume that  $f(z) \neq 0$  only for  $z > 0$ , where our result for  $G$  applies. Then the integrand is 0 for  $z < 0$  (because  $f = 0$  there) and also for  $z > x$  (because  $G = 0$  there). Thus

$$\begin{aligned} y(x) &= \int_0^x G(x, z) f(z) dz \\ &= \int_0^x \frac{y_1(z)f(z)}{W(z)} dz y_2(x) - \int_0^x \frac{y_2(z)f(z)}{W(z)} dz y_1(x). \end{aligned}$$

This is exactly the same solution that is found in differential equations textbooks by making the ansatz

$$y(x) = u_1(x)y_1(x) + u_2(x)y_2(x)$$

and deriving a system of first-order differential equations for  $u_1$  and  $u_2$ . That method is called “variation of parameters”. Writing the variation-of-parameters solution in terms of the Green function  $G$  shows in a precise and clear way how the solution  $y$  depends on the nonhomogeneous term  $f$  as  $f$  is varied. That formula is a useful starting point for many further investigations.

**Example 3:** *Inhomogeneous boundary data.* Consider the problem

PDE: 
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

BC: 
$$u(x, 0) = \delta(x - z).$$

Its solution is

$$G(x - z, y) \equiv \frac{1}{\pi} \frac{y}{(x - z)^2 + y^2},$$

the Green function that we constructed for Laplace’s equation in the upper half plane. Therefore, the general solution of Laplace’s equation in the upper half plane, with arbitrary initial data

$$u(x, 0) = f(x),$$

is

$$u(x, y) = \int_{-\infty}^{\infty} dz G(x - z, y) f(z).$$

Similarly, the Green function

$$G(t, x - z) = \frac{1}{\sqrt{4\pi t}} e^{-(x-z)^2/4t}.$$

that we found for the heat equation solves the heat equation with initial data

$$u(0, x) = \delta(x - z).$$

And so on, for any linear problem with nonhomogeneous data.

### DELTA FUNCTIONS AND FOURIER TRANSFORMS

Formally, the Fourier transform of a delta function is a complex exponential function, since

$$\int_{-\infty}^{\infty} \delta(x - z) e^{-ikx} dx = e^{-ikz}.$$

According to the Fourier inversion formula, therefore, we should have

$$\begin{aligned} \delta(x - z) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikz} e^{ikx} dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-z)} dk. \end{aligned}$$

This is a very useful formula! Here is another way of seeing what it means and why it is true:

Recall that

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk, \\ \hat{f}(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(z) e^{-ikz} dz. \end{aligned}$$

Let us substitute the second formula into the first:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ik(x-z)} f(z) dz dk.$$

Of course, this equation is useless for computing  $f(x)$ , since it just goes in a circle; its significance lies elsewhere. If we're willing to play fast and loose with the order of the integrations, we can write it

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} e^{ik(x-z)} dk \right] f(z) dz,$$

which says precisely that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-z)} dk$$

satisfies the defining property of  $\delta(x-z)$ . Our punishment for playing fast and loose is that this integral does not converge (in the usual sense), and there is no function  $\delta$  with the desired property. Nevertheless, both the integral and the object  $\delta$  itself can be given a rigorous meaning in the modern theory of *distributions*; crudely speaking, they both make perfect sense as long as you keep them inside other integrals (multiplied by continuous functions) and do not try to evaluate them at a point to get a number.

What would happen if we tried this same trick with the Fourier series formulas? Let's consider the sine series,

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx,$$

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(z) \sin nz dz.$$

This gives

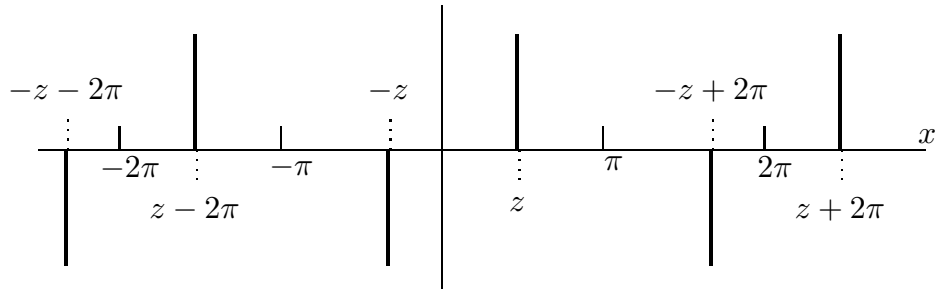
$$f(x) = \frac{2}{\pi} \int_0^{\pi} \left[ \sum_{n=1}^{\infty} \sin nx \sin nz \right] f(z) dz. \quad (\dagger)$$

Does this entitle us to say that

$$\delta(x-z) = \frac{2}{\pi} \sum_{n=1}^{\infty} \sin nx \sin nz? \quad (\ddagger)$$

Yes and no. In  $(\dagger)$  the variables  $x$  and  $z$  are confined to the interval  $[0, \pi]$ .  $(\ddagger)$  is a valid representation of the delta function when applied to functions whose domain is  $[0, \pi]$ . If we applied it to a function on a larger domain, it would act like the odd, periodic extension of  $\delta(x-z)$ , as is always the case with Fourier sine series:

$$\frac{2}{\pi} \sum_{n=1}^{\infty} \sin nx \sin nz = \sum_{M=-\infty}^{\infty} [\delta(x-z+2\pi M) - \delta(x+z+2\pi M)].$$



THE POISSON SUMMATION FORMULA

NOTE: This is *not* what Haberman calls “Poisson formula” in Exercise 2.5.4 and p. 433.

Let’s repeat the foregoing discussion for the case of the full Fourier series on the interval  $(-L, L)$ :

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\pi n x/L}, \quad c_n = \frac{1}{2L} \int_{-L}^L e^{-i\pi n y/L} f(y) dy$$

leads to

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} \frac{1}{2L} \int_{-L}^L e^{-i\pi y/L} f(y) dy e^{i\pi n x/L} \\ &= \int_{-L}^L \left[ \frac{1}{2L} \sum_{n=-\infty}^{\infty} e^{i\pi n(x-y)/L} \right] f(y) dy. \end{aligned}$$

Therefore,

$$\frac{1}{2L} \sum_{n=-\infty}^{\infty} e^{i\pi n(x-y)/L} = \delta(x-y) \quad \text{for } x \text{ and } y \text{ in } (-L, L).$$

Now consider  $y = 0$  (for simplicity). For  $x$  outside  $(-L, L)$ , the sum must equal the  $2L$ -periodic extension of  $\delta(x)$ :

$$\frac{1}{2L} \sum_{n=-\infty}^{\infty} e^{i\pi n x/L} = \sum_{M=-\infty}^{\infty} \delta(x - 2LM). \quad (\ddagger)$$

Let  $f(x)$  be a continuous function on  $(-\infty, \infty)$  whose Fourier transform is also continuous. Multiply both sides of  $(\ddagger)$  by  $f(x)$  and integrate:

$$\frac{\sqrt{2\pi}}{2L} \sum_{n=-\infty}^{\infty} \hat{f}\left(-\frac{\pi n}{L}\right) = \sum_{M=-\infty}^{\infty} f(2LM).$$

Redefine  $n$  as  $-n$  and simplify:

$$\sqrt{\frac{\pi}{2}} \frac{1}{L} \sum_{n=-\infty}^{\infty} \hat{f}\left(+\frac{\pi n}{L}\right) = \sum_{M=-\infty}^{\infty} f(2LM).$$

This *Poisson summation formula* says that the sum of a function over an equally spaced grid of points equals the sum of its Fourier transform over a certain

other equally spaced grid of points. The most symmetrical version comes from choosing  $L = \sqrt{\frac{\pi}{2}}$

$$\sum_{n=-\infty}^{\infty} \hat{f}(\sqrt{2\pi} n) = \sum_{M=-\infty}^{\infty} f(\sqrt{2\pi} M).$$

However, the most frequently used version, and probably the easiest to remember, comes from taking  $L = \frac{1}{2}$ : Starting with a numerical sum

$$\sum_{M=-\infty}^{\infty} f(M),$$

one can replace it by

$$\sqrt{2\pi} \sum_{n=-\infty}^{\infty} \hat{f}(2\pi n),$$

which is

$$\sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-2\pi i n x} dx$$

(and the minus sign in the exponent is unnecessary).



## Additional Topics on Green Functions

### A GREEN FUNCTION FOR THE WAVE EQUATION

It is relatively difficult to work the Fourier-transform solution of the wave equation into a Green-function form, because the integrals are poorly convergent. However, we already have a Green-function solution of the initial-value problem for the wave equation: it is d'Alembert's solution! Recall that (for  $c = 1$ ,  $f(x) \equiv u(0, x)$ ,  $g(x) \equiv \frac{\partial u}{\partial t}(0, x)$ ) the solution is

$$u(t, x) = \frac{1}{2}[f(x+t) + f(x-t)] + \frac{1}{2} \int_{x-t}^{x+t} g(z) dz. \quad (1)$$

For simplicity consider only the case  $t > 0$ . Then (1) can be written

$$\begin{aligned} u(t, x) = & \frac{1}{2} \int_{-\infty}^{\infty} dz f(z) [\delta(z-x-t) + \delta(z-x+t)] \\ & + \frac{1}{2} \int_{-\infty}^{\infty} dz g(z) [h(z-x+t) - h(z-x-t)], \end{aligned} \quad (2)$$

where  $h$  is the unit step function; recall that it satisfies

$$\delta(w) = \frac{dh(w)}{dw}.$$

Now define

$$G(t, x, z) \equiv \frac{1}{2}[h(z-x+t) - h(z-x-t)],$$

so that

$$\frac{\partial G}{\partial t}(t, x, z) = \frac{1}{2}[\delta(z-x+t) + \delta(z-x-t)].$$

Then (2) can be rewritten as

$$u(t, x) = \int_{-\infty}^{\infty} \frac{\partial G}{\partial t}(t, x, z) u(0, z) dz + \int_{-\infty}^{\infty} G(t, x, z) \frac{\partial u}{\partial t}(0, z) dz.$$

(Although we assumed  $t > 0$ , this formula also holds for  $t < 0$ .)

This particular kind of combination of *boundary values and derivatives of the solution and a Green function* is quite characteristic of boundary-value problems for second-order equations. We'll see it again in connection with Laplace's equation.

For a variety of historical and practical reasons, this course concentrates on *homogeneous* linear PDEs and their (nonhomogeneous) boundary-value problems. From a Green-function point of view, however, nonhomogeneous differential equations are actually more fundamental. We will look briefly at two of these.

### The Green function for the heat equation with source

Recall that the solution of the initial-value problem for the homogeneous heat equation is

$$u(t, x) = \int_{-\infty}^{\infty} H(t, x, y) f(y) dy \quad (f(x) \equiv u(0, x)),$$

where

$$H(t, x, y) = \frac{1}{\sqrt{4\pi t}} e^{-(x-y)^2/4t}.$$

$H$  could be defined as the solution of the initial-value problem

$$\frac{\partial H}{\partial t} = \frac{\partial^2 H}{\partial x^2}, \quad H(0, x, y) = \delta(x - y). \quad (4)$$

We are now interested in the nonhomogeneous heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + j(t, x) \quad (\text{for } t > 0), \quad u(0, x) = 0 \quad (5)$$

(where we've imposed the homogeneous initial condition to make the solution unique). In view of our experience with ODEs we might expect the solution to be of the form

$$u(t, x) = \int_{-\infty}^{\infty} dy \int_0^{\infty} ds G(t, x; s, y) j(s, y), \quad (6)$$

where  $G$  satisfies

$$\frac{\partial G}{\partial t} - \frac{\partial^2 G}{\partial x^2} = \delta(t - s)\delta(x - y), \quad G(0, x, s, y) = 0 \quad (7)$$

(i.e., the temperature response to a *point source* of heat at position  $y$  and time  $s$ .) The surprising fact is that  $G$  turns out to be essentially the same thing as  $H$ .

To see that, consider

$$u(t, x) \equiv \int_{-\infty}^{\infty} dy \int_0^t ds H(t - s, x, y) j(s, y).$$

It can be proved that differentiation “under the integral sign” is legitimate here, so let’s just calculate

$$\begin{aligned}\frac{\partial u}{\partial t} &= \int_{-\infty}^{\infty} dy H(0, x, y)j(t, y) + \int_{-\infty}^{\infty} dy \int_0^t ds \frac{\partial H}{\partial t}(t - s, x, y)j(s, y), \\ \frac{\partial^2 u}{\partial x^2} &= \int_{-\infty}^{\infty} dy \int_0^t ds \frac{\partial^2 H}{\partial x^2}(t - s, x, y)j(s, y).\end{aligned}$$

Now use (4) to evaluate the first term in  $\frac{\partial u}{\partial t}$  and to observe that the other term cancels  $\frac{\partial^2 u}{\partial x^2}$  when we construct

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = \int_{-\infty}^{\infty} dy \delta(x - y)j(t, y) = j(t, x).$$

Also, we have  $u(0, x) = 0$ . So our  $u$  solves the problem (5). In other words, the solutions of (5) is (6) with

$$G(t, x; s, y) = \begin{cases} H(t - s, x, y) & \text{if } s \leq t, \\ 0 & \text{if } s > t. \end{cases}$$

Put the other way around: The Green function that solves the initial-value problem for the homogeneous heat equation is

$$H(t, x, y) = G(t, x; 0, y),$$

where  $G$  is the Green function that solves the nonhomogeneous heat equation with homogeneous initial data (and is defined by (7)). This connection between nonhomogeneous and homogeneous Green functions is called *Duhamel’s principle* (specifically, for the heat equation, and more loosely, for analogous more general situations).

The previous result for the wave equation is another instance of this principle: It can be shown that

$$\begin{aligned}G_{\text{ret}}(t, x; s, z) &\equiv G(t - s, x, z)h(t - s) \\ &= \frac{1}{2}[h(z - x + t - s) - h(z - x - t + s)]h(t - s) \\ &= \frac{1}{2}h(t - x)h(t + x)\end{aligned}$$

is a Green function for the nonhomogeneous wave equation, in the sense that

$$u(t, x) = \int_{-\infty}^{\infty} dy \int_{-\infty}^t ds G_{\text{ret}}(t, x; s, y)f(s, y)$$

satisfies

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = f(t, x).$$

(Here the  $G(t-s, x, z)$  is the one previously constructed for the wave equation.) The subscript “ret” stands for *retarded*. It means that the effects of the source  $f$  show up only *later in time*. (Pictorially, a point source at  $(s, y)$  emits a wave into the forward-pointing space-time cone of points  $(t, x)$  with its vertex at the source. Elsewhere  $G_{\text{ret}} = 0$ .) Because the wave equation is second-order and time-symmetric, there are infinitely many other Green functions, corresponding to different initial conditions. In particular, there is an *advanced* Green function that absorbs everything and leaves the space empty of waves at later times. For thermodynamic reasons the retarded solution is the relevant one in most applications. (You do not often turn on a flashlight with an incoming wave already focused on it.)

We shall soon see the Duhamel principle at work for Laplace’s equation, too.

### Coulomb fields

The nonhomogeneous version of Laplace’s equation,

$$-\nabla^2 u = j(\mathbf{x}),$$

is called *Poisson’s equation*. The corresponding Green function, satisfying

$$-\nabla_{(\mathbf{x})}^2 G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}),$$

has the physical interpretation of the electrostatic field at  $\mathbf{x}$  created by a point charge at  $\mathbf{y}$ . (The subscript “ $(\mathbf{x})$ ” tells us which variable the operator acts upon.) In dimension 3, with  $r \equiv \|\mathbf{x} - \mathbf{y}\|$ , it is well known to be

$$G(\mathbf{x}, \mathbf{y}) = G(\mathbf{x} - \mathbf{y}) = \frac{1}{4\pi r}$$

(times constants that depend on the system of electrical units being used). In general dimension  $n$  (greater than 2, a special case) this becomes

$$G(\mathbf{x}, \mathbf{y}) = \frac{C}{r^{n-2}},$$

where  $[(n-2)C]^{-1}$  is the “surface area” of the unit  $(n-1)$ -sphere. For  $n = 2$  the formula is

$$G(\mathbf{x}, \mathbf{y}) = -\frac{\ln r}{2\pi} = -\frac{\ln r^2}{4\pi}.$$

**Sketch of proof:** For  $r \neq 0$  one has in  $n$ -dimensional spherical coordinates

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{n-1}{r} \frac{\partial}{\partial r} + \frac{\partial}{\partial(\text{angles})},$$

so  $\nabla^2 r^{2-n} = 0$ , as required. Now the hard part is showing that the function has the delta behavior at the origin. Let  $B_\epsilon$  be the ball of radius  $\epsilon$  centered at  $\mathbf{y}$ , and let  $S_\epsilon$  be its boundary (a sphere of radius  $\epsilon$ ). If we trust that Gauss's theorem continues to hold if delta functions in the derivatives are taken into account, then

$$\begin{aligned} \int_{B_\epsilon} \nabla^2 G d^n z &= \int_{S_\epsilon} \hat{\mathbf{n}} \cdot \nabla G d^{n-1} S = \int_{S_\epsilon} \frac{\partial G}{\partial r} d^{n-1} S \\ &= C(2-n) \int_{S_\epsilon} \epsilon^{1-n} \epsilon^{n-1} d(\text{angles}) \\ &= -(n-2)C \times (\text{area of sphere of unit radius}) = -1. \end{aligned}$$

Thus the singularity at the origin has the correct normalization. To make a real proof one should do two things: (1) We really need to show, not just that  $\int \nabla^2 G = -1$ , but that  $\int \nabla^2 G(\mathbf{z}) f(\mathbf{z}) d^n z = -f(\mathbf{0})$  for all smooth functions  $f$ . This is not much harder than the calculation just shown: Either use ‘‘Green’s symmetric identity’’ (reviewed in a later subsection), or expand  $f$  in a power series. All the unwanted terms will go to zero as  $\epsilon \rightarrow 0$ . (2) Strictly speaking, the action of  $\nabla^2$  on  $G$  is *defined* by integration by parts (in the whole space):

$$\int_{\mathbf{R}^n} \nabla^2 G(\mathbf{z}) f(\mathbf{z}) d^n z \equiv \int_{\mathbf{R}^n} G(\mathbf{z}) \nabla^2 f(\mathbf{z}) d^n z,$$

where  $f$  is assumed to vanish at infinity. Now apply Gauss’s theorem to the *outside* of  $S_\epsilon$ , where we know it is valid, to show that this integral equals  $-f(\mathbf{0})$ .

## THE METHOD OF IMAGES

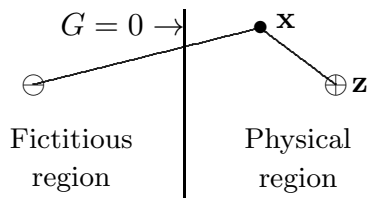
The image method is a generalization of the solution of the wave equation by even and odd periodic extensions, except that this time we extend the Green function instead of the initial data. It is simplest and most intuitive for nonhomogeneous equations, but we’ll see that it can easily be extended to homogeneous equations with initial data. It is easy to treat the Poisson and heat equations simultaneously and in parallel.

### A single Dirichlet or Neumann boundary

Consider the Poisson Green-function equation

$$-\nabla^2 G(\mathbf{x}, \mathbf{z}) = \delta(\mathbf{x} - \mathbf{z})$$

for  $\mathbf{x}$  and  $\mathbf{z}$  in a half-space, with the PDE’s solutions, and hence  $G$ , required to vanish on its boundary (a ‘‘perfectly conducting plane’’ in physics terminology). Start with the Coulomb potential of the source at  $\mathbf{z}$  (a positive charge). If we also place a charge of opposite sign in the mirror-image position opposite the charge at  $\mathbf{z}$ , then *its* Coulomb potential satisfies  $\nabla^2 u = 0$  in the physical region (so it doesn’t mess up the property  $-\nabla^2 G = \delta$ ), and on the boundary its Coulomb field (gradient) precisely cancels the Coulomb field of the original charge. Success!



To write this construction down algebraically, we need to choose a good coordinate system. Consider  $n = 2$  for simplicity; without loss of generality put  $\mathbf{z}$  on the  $x$  axis, so  $\mathbf{z} = (z, 0)$ ; write  $\mathbf{x} = (x, y)$  with the boundary along the  $y$  axis,  $x = 0$ . Then our Green function is

$$G(\mathbf{x}, \mathbf{z}) = -\frac{1}{4\pi} \ln[(x - z)^2 + y^2] + \frac{1}{4\pi} \ln[(x + z)^2 + y^2], \quad (8)$$

because  $(x - z)^2 + y^2$  is the square of the distance from  $\mathbf{x}$  to the positive charge and  $(x + z)^2 + y^2$  is the square of the distance to the fictitious negative charge. Notice that  $G(\mathbf{x}, \mathbf{z}) \neq G(\mathbf{x} - \mathbf{z})$  in this problem, unlike the Coulomb potential and all the other simple Green functions we have seen for *translation-invariant* problems. (This problem is *not* invariant under translations, because the boundary is fixed at  $x = 0$ .)

The extension of this construction to higher dimensions is easy, but alphabetically inconvenient if you insist on using scalar variables. It would be better to introduce a notation for the components of vectors parallel and perpendicular to the boundary.

Similarly, the Green function for the heat equation on a half-line with  $u(0, x) = 0$  is

$$H(t, x, y) - H(t, x, -y) = \frac{1}{\sqrt{4\pi t}} \left[ e^{-(x-y)^2/4t} - e^{-(x+y)^2/4t} \right]. \quad (9)$$

This can be shown equal to the Fourier solution

$$\frac{2}{\pi} \int_0^\infty \sin(kx) \sin(ky) e^{-k^2 t} dk. \quad (10)$$

Function (9) is the Green function for the nonhomogeneous heat equation with the source at time  $s = 0$  (from which the general case can be obtained by the substitution  $t \leftarrow t - s$ ), but by Duhamel's principle it is also the Green function for the homogeneous heat equation with initial data given at  $t = 0$ , and it is in that role that we have previously encountered (10).

To solve a *Neumann* problem at  $x = 0$  ( $\frac{\partial u}{\partial x} = 0$  there, so  $\frac{\partial G}{\partial x} = 0$ ), we simply *add* the contribution of the image source instead of subtracting it. This produces a solution that is *even* (rather than odd) under reflection through the boundary plane, and hence its *normal derivative* (rather than the function itself) vanishes on the plane.

## The periodic case

Suppose we are interested in the initial-value problem for the heat equation on a ring with coordinate  $x$ ,  $-\pi < x \leq \pi$ . We know that the relevant Green function is

$$K(t, x, y) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{in(x-y)} e^{-n^2 t} \quad (11)$$

— from substituting

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iny} f(y) dy \quad \text{into} \quad u(t, x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} e^{-n^2 t}.$$

But another way to get such a Green function is to start from the one for the whole line,  $H$ , and add copies of it spaced out periodically:

$$K(t, x, y) = \sum_{M=-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{-(x-y-2\pi M)^2/4t}. \quad (12)$$

Each term of (12) (and hence the whole sum) satisfies the heat equation for  $t > 0$ . As  $t \rightarrow 0$  the term with  $M = 0$  approaches  $\delta(x - y)$  as needed, and all the other terms approach 0 if  $x$  and  $y$  are in the basic interval  $(-\pi, \pi)$ . Finally, the function is periodic,  $K(t, x + 2\pi, y) = K(t, x, y)$ , as desired.

The functions (11) and (12) are equal, although this is not obvious by inspection. Neither sum can be evaluated in closed form in terms of elementary functions. From a numerical point of view they are useful in complementary domains, because the sum in (12) converges very fast when  $t$  is small, whereas the one in (11) converges best when  $t$  is large.

The equality of (11) and (12) is an instance of the *Poisson summation formula*. This is most easily seen when  $x = y$ , so that the equality is

$$\sum_{M=-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{-(2\pi M)^2/4t} = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} e^{-n^2 t}. \quad (13)$$

Since

$$H(t, z) = \frac{1}{\sqrt{4\pi t}} e^{-z^2/4t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikz} \hat{H}(t, k) dk$$

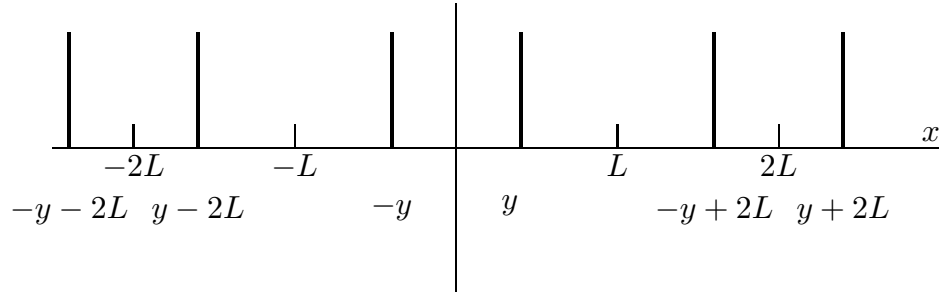
where  $\hat{H}(t, k) = \frac{1}{\sqrt{2\pi}} e^{-k^2 t}$ , (13) is the Poisson relation

$$\sum_{M=-\infty}^{\infty} H(t, 2\pi M) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \hat{H}(t, n).$$

## Finite intervals

On the interval  $(0, L)$  with Dirichlet boundary conditions at both ends, or Neumann boundary conditions at both ends, we can get the Green function by combining the two previous ideas.

In the Neumann case, reflect the source (at  $y$ ) through both ends to get “image charges” at  $-y$  and  $2L - y$ . Continue this process indefinitely in both directions to get an infinite sequence of images that build up the needed *even periodic extension* of the delta functions and hence of the Green function and, ultimately, of the solution of the PDE problem.



In the Dirichlet case the first two images are *negative*, and thereafter they alternate in sign so as to build up the odd periodic extensions. (Compare the end of the previous section, where the corresponding linear combination of delta functions was sketched.)

## APPLICATION OF GREEN'S IDENTITY

If  $V$  is a region in space bounded by a surface  $S$ , and  $u$  and  $v$  are two functions, then Gauss's theorem applied to the vector field  $u\nabla v - v\nabla u$  implies

$$\int_S (u\nabla v - v\nabla u) \cdot \hat{\mathbf{n}} dS = \int_V (u\nabla^2 v - v\nabla^2 u) d^3x. \quad (14)$$

Here  $\hat{\mathbf{n}}$  is the outward unit normal vector on  $S$ , so  $\hat{\mathbf{n}} \cdot \nabla u$  (for example) is the outward *normal derivative* of  $u$ , the quantity that appears in Neumann boundary conditions. In the simple regions we have studied so far, it was always possible to write the normal derivative as  $(\pm)$  the partial derivative in a coordinate that is constant on a portion of the boundary. Formula (14) makes sense and holds true in any dimension, not just 3. It is called *Green's symmetric identity* or *Green's second identity*.



Green's identity has many applications to PDEs, of which we can demonstrate only one of the simplest. Suppose that  $G(\mathbf{x}, \mathbf{y})$  is the Green function that solves the *homogeneous Dirichlet problem* for the *Poisson equation* in  $V$ :

$$-\nabla_{(\mathbf{x})}^2 G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}) \quad \text{for } \mathbf{x} \in V, \quad G(\mathbf{x}, \mathbf{y}) = 0 \quad \text{for } \mathbf{x} \in S.$$

Let  $u(\mathbf{x})$  be any solution of Laplace's equation in  $V$ :  $\nabla^2 u = 0$ . Apply (14) with  $G$  in the role of  $v$ :

$$\int_V [u(\mathbf{x})\nabla^2 G(\mathbf{x}, \mathbf{y}) - G(\mathbf{x}, \mathbf{y})\nabla^2 u(\mathbf{x})] d^3x = \int_S [u(\mathbf{x})\nabla G(\mathbf{x}, \mathbf{y}) - G(\mathbf{x}, \mathbf{y})\nabla u(\mathbf{x})] \cdot \hat{\mathbf{n}} dS.$$

By the conditions defining  $G$  and  $u$ , this reduces to

$$u(\mathbf{y}) = \int_V u(\mathbf{x})\delta(\mathbf{x} - \mathbf{y}) d^3x = - \int_S \hat{\mathbf{n}} \cdot \nabla_{(\mathbf{x})} G(\mathbf{x}, \mathbf{y})u(\mathbf{x}) dS \equiv \int_S g(\mathbf{y}, \mathbf{x})u(\mathbf{x}) dS.$$

This formula expresses  $u$  in terms of its Dirichlet data on  $S$ . It therefore solves the *nonhomogeneous Dirichlet problem* for *Laplace's equation* in  $V$ . This is the version of Duhamel's principle that applies to this situation.

For example, let  $V$  be the upper half plane. By the method of images ((8) above with the coordinates turned around), the Green function is

$$G(\mathbf{x}, \mathbf{y}) = -\frac{1}{4\pi} \ln[(x_1 - y_1)^2 + (x_2 - y_2)^2] + \frac{1}{4\pi} \ln[(x_1 - y_1)^2 + (x_2 + y_2)^2].$$

(Here  $\mathbf{y} = (y_1, y_2)$ , etc., and the image charge is at  $(y_1, -y_2)$ .) To get  $g(\mathbf{y}, \mathbf{x}) \equiv -\hat{\mathbf{n}} \cdot \nabla_{(\mathbf{x})} G(\mathbf{x}, \mathbf{y})$  we need to differentiate with respect to  $-x_2$  (since the outward direction is *down*) and evaluate it at  $x_2 = 0$  (the boundary  $S$ ). This gives

$$-\frac{1}{2\pi} [(x_1 - y_1)^2 + y_2^2]^{-1} (-y_2) + \frac{1}{2\pi} [(x_1 - y_1)^2 + y_2^2]^{-1} (+y_2) = \frac{y_2}{\pi} [(x_1 - y_1)^2 + y_2^2]^{-1}.$$

Reverting to our usual notation ( $x_1 \rightarrow z$ ,  $y_2 \rightarrow y$ ,  $y_1 \rightarrow z$ ) we get

$$g(x, z; y) = \frac{1}{\pi} \frac{y}{(x - z)^2 + y^2},$$

our old Green function for this problem!

## Sturm–Liouville Problems

### MORE GENERAL EIGENVALUE PROBLEMS

So far all of our example PDEs have led to separated equations of the form  $X'' + \omega^2 X = 0$ , with standard Dirichlet or Neumann boundary conditions. Not surprisingly, more complicated equations often come up in practical problems. For example, if the medium in a heat or wave problem is spatially inhomogeneous,\* the relevant equation may look like

$$X'' - V(x)X = -\omega^2 X$$

for some function  $V$ , or even

$$a(x)X'' + b(x)X' + c(x)X = -\omega^2 X.$$

Also, if the boundary in a problem is a circle, cylinder, or sphere, the solution of the problem is simplified by converting to polar, cylindrical, or spherical coordinates, so that the boundary is a surface of constant radial coordinate. This simplification of the boundary conditions is bought at the cost of complicating the differential equation itself: we again have to deal with ODEs with nonconstant coefficients, such as

$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} - \frac{n^2}{r^2} R = -\omega^2 R.$$

The good news is that many of the properties of Fourier series carry over to these more general situations. As before, we can consider the *eigenvalue problem* defined by such an equation together with appropriate boundary conditions: Find all functions that satisfy the ODE (for *any* value of  $\omega$ ) and also satisfy the boundary conditions. And it is still true (under certain conditions) that the set of all eigenfunctions is *complete*: Any reasonably well-behaved function can be expanded as an infinite series where each term is proportional to one of the eigenfunctions. This is what allows arbitrary data functions in the original PDE to be matched to a sum of separated solutions! Also, the eigenfunctions are *orthogonal* to each other; this leads to a simple formula for the coefficients in the eigenfunction expansion, and also to a Parseval formula relating the norm of the function to the sum of the squares of the coefficients.

---

\* That is, the density, etc., vary from point to point. This is not the same as “nonhomogeneous” in the sense of the general theory of linear differential equations.

## ORTHONORMAL BASES

Consider an interval  $[a, b]$  and the real-valued (or complex-valued) functions defined on it. A sequence of functions  $\{\phi_n(x)\}$  is called *orthogonal* if

$$\int_a^b \phi_n(x)^* \phi_m(x) dx = 0 \quad \text{whenever } m \neq n.$$

It is called *orthonormal* if, in addition,

$$\int_a^b |\phi_n(x)|^2 dx = 1.$$

This normalization condition is merely a convenience; the important thing is the orthogonality. (If we are lucky enough to have an orthogonal set, we can always convert it to an orthonormal set by dividing each function by the square root of its normalization integral:

$$\psi_n(x) \equiv \frac{\phi_n(x)}{\sqrt{\int_a^b |\phi_n(z)|^2 dz}} \Rightarrow \int_a^b |\psi_n(x)|^2 dx = 1.$$

However, in certain cases this may make the formula for  $\psi_n$  more complicated, so that the redefinition is hardly worth the effort. A prime example is the eigenfunctions in the Fourier sine series:

$$\phi_n(x) \equiv \sin nx \Rightarrow \int_0^\pi |\phi_n(x)|^2 dx = \frac{\pi}{2};$$

therefore,

$$\psi_n(x) \equiv \sqrt{\frac{2}{\pi}} \sin nx$$

are the elements of the orthonormal basis. (This is the kind of normalization often used for the Fourier sine *transform*, as we have seen.) A good case can be made, however, that normalizing the eigenfunctions is more of a nuisance than a help in this case; most people prefer to put the entire  $2/\pi$  in one place rather than put half of it in the Fourier series and half in the coefficient formula.

Now let  $f(x)$  be an arbitrary (nice) function on  $[a, b]$ . If  $f$  has an expansion as a linear combination of the  $\phi$ 's,

$$f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x),$$

then

$$\int_a^b \phi_m(x)^* f(x) dx = \sum_{n=1}^{\infty} c_n \int_a^b \phi_m(x)^* \phi_n(x) dx = c_m \int_a^b |\phi_m(x)|^2 dx$$

by orthogonality. If the set is orthonormal, this just says

$$\boxed{c_m = \int_a^b \phi_m(x)^* f(x) dx.} \quad (\P)$$

(In the rest of this discussion, I shall assume that the orthogonal set is orthonormal. This greatly simplifies the formulas of the general theory, even while possibly complicating the expressions for the eigenfunctions in any particular case.)

It can easily be shown that

$$\boxed{\int_a^b |f(x)|^2 dx = \sum_{n=1}^{\infty} |c_n|^2.}$$

This is the *Parseval equation* associated to this orthonormal set. Furthermore, if  $f$  is *not* of the form  $\sum_{n=1}^{\infty} c_n \phi_n(x)$ , then

$$(1) \quad \sum_{n=1}^{\infty} |c_n|^2 < \int_a^b |f(x)|^2 dx$$

(called *Bessel's inequality*), and (2) the *best approximation* to  $f(x)$  of the form  $\sum c_n \phi_n(x)$  is the one where the coefficients are computed by formula  $(\P)$ . These last two statements remain true when  $\{\phi_n\}$  is a *finite* set — in which case, obviously, the probability that a given  $f$  will not be exactly a linear combination of the  $\phi$ 's is greatly increased. (The precise meaning of (2) is that the choice  $(\P)$  of the  $c_n$  minimizes the integral

$$\int_a^b \left| f(x) - \sum_{n=1}^{\infty} c_n \phi_n(x) \right|^2 dx.$$

That is, we are talking about *least squares approximation*. It is understood in this discussion that  $f$  itself is square-integrable on  $[a, b]$ .)

Now suppose that every square-integrable  $f$  is the limit of a series  $\sum_{n=1}^{\infty} c_n \phi_n$ . (This series is supposed to converge “in the mean” — that is, the least-squares integral

$$\int_a^b \left| f(x) - \sum_{n=1}^M c_n \phi_n(x) \right|^2 dx$$

for a partial sum approaches 0 as  $M \rightarrow \infty$ .) Then  $\{\phi_n\}$  is called a *complete set* or an *orthonormal basis*. This is the analogue of the mean convergence theorem for Fourier series. Under certain conditions there may also be pointwise or uniform convergence theorems, but these depend more on the special properties of the particular functions  $\phi$  being considered.

So far this is just a definition, not a theorem. To guarantee that our orthonormal functions form a basis, we have to know where they came from. The miracle of the subject is that the eigenfunctions that arise from variable-separation problems *do* form orthonormal bases:

### STURM-LIOUVILLE THEORY

**Theorem:** Suppose that the ODE that arises from some separation of variables is

$$\mathcal{L}[X] = -\omega^2 r(x)X \quad \text{on } (0, L), \quad (\ddagger)$$

where  $\mathcal{L}$  is an abbreviation for a second-order *linear differential operator*

$$\mathcal{L}[X] \equiv a(x)X'' + b(x)X' + c(x)X,$$

$a$ ,  $b$ ,  $c$ , and  $r$  are continuous on  $[0, L]$ , and  $a(x) > 0$  and  $r(x) > 0$  on  $[0, L]$ . Suppose further that

$$\int_0^L (\mathcal{L}[u](x))^* v(x) dx = \int_0^L u(x)^* (\mathcal{L}[v](x)) dx \quad (\dagger)$$

for all functions  $u$  and  $v$  satisfying the boundary conditions of the problem. (In terms of the inner product in  $L^2$ , this condition is just  $\langle \mathcal{L}u, v \rangle = \langle u, \mathcal{L}v \rangle$ .) Then:

- (1) All the eigenvalues  $\omega^2$  are real (but possibly negative).
- (2) The eigenfunctions corresponding to different  $\omega$ 's are orthogonal with respect to the *weight function*  $r(x)$ :

$$\int_0^L \phi_n(x)^* \phi_m(x) r(x) dx = 0 \quad \text{if } n \neq m.$$

(Everything said previously about orthonormality can be generalized to the case of a nontrivial positive weight function. Here we are really dealing with a *new* inner product,

$$\langle u, v \rangle_r \equiv \int_0^L u(x)^* v(x) r(x) dx,$$

and also a new operator,  $\mathcal{A}[X] \equiv \mathcal{L}[X]/r$ , so that the differential equation  $(\ddagger)$  is the eigenvalue equation  $\mathcal{A}[X] = -\omega^2 X$ . The two factors of  $r$  cancel in  $(\dagger)$ .)

- (3) The eigenfunctions are complete. (This implies that the corresponding PDE can be solved for arbitrary boundary data, in precise analogy to Fourier series problems!)

The proof that a given  $\mathcal{L}$  satisfies (†) (or doesn't satisfy it, as the case may be) involves just integrating by parts twice. (Setting  $v$  equal to  $u$  in the intermediate step of this calculation gives, as a bonus, a proof of part (7) in the continuation of this theorem below. You are invited to fill in the details.) It turns out that (†) will be satisfied if  $\mathcal{L}$  has the form

$$\frac{d}{dx}p(x)\frac{d}{dx} + q(x)$$

(with  $p$  and  $q$  real-valued and well-behaved) and the boundary conditions are of the type

$$\alpha X'(0) - \beta X(0) = 0, \quad \gamma X'(L) + \delta X(L) = 0$$

with  $\alpha$ , etc., real.\* Such an eigenvalue problem is called a *regular Sturm–Liouville problem*.

The proof of the conclusions (1) and (2) of the theorem is quite simple and is a generalization of the proof of the corresponding theorem for eigenvalues and eigenvectors of a *symmetric matrix* (which is proved in many physics courses and linear algebra courses). Part (3) is harder to prove, like the convergence theorems for Fourier series (which are a special case of it).

#### EXAMPLE: CONVECTIVE BOUNDARY CONDITION

The simplest nontrivial example of a Sturm–Liouville problem (“nontrivial” in the sense that it gives something other than a Fourier series) is the usual spatially homogeneous heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (0 < x < L, \quad 0 < t < \infty),$$

with boundary conditions such as

$$u(0, t) = 0, \quad \frac{\partial u}{\partial x}(L, t) + \beta u(L, t) = 0$$

and initial data

$$u(x, 0) = f(x).$$

---

\* The reason for the minus sign in the first equation is to make true “property (7)” stated below.

In a realistic problem, the zeros in the BC would be replaced by constants; as usual, we would take care of that complication by subtracting off a steady-state solution. Physically, the constant value of  $\frac{\partial u}{\partial x}(L, t) + \beta u(L, t)$  is proportional to the temperature of the air (or other fluid medium) to which the right-hand endpoint of the bar is exposed; heat is lost through that end by convection, according to “Newton’s law of cooling”. Mathematically, such a BC is called a *Robin boundary condition*, as opposed to Dirichlet or Neumann.

The separation of variables proceeds just as in the more standard heat problems, up to the point

$$T(t) = e^{-\omega^2 t}, \quad X(x) = \sin \omega x.$$

To get the sine I used the boundary condition  $X(0) = 0$ . The other BC is

$$X'(L) + \beta X(L) = 0,$$

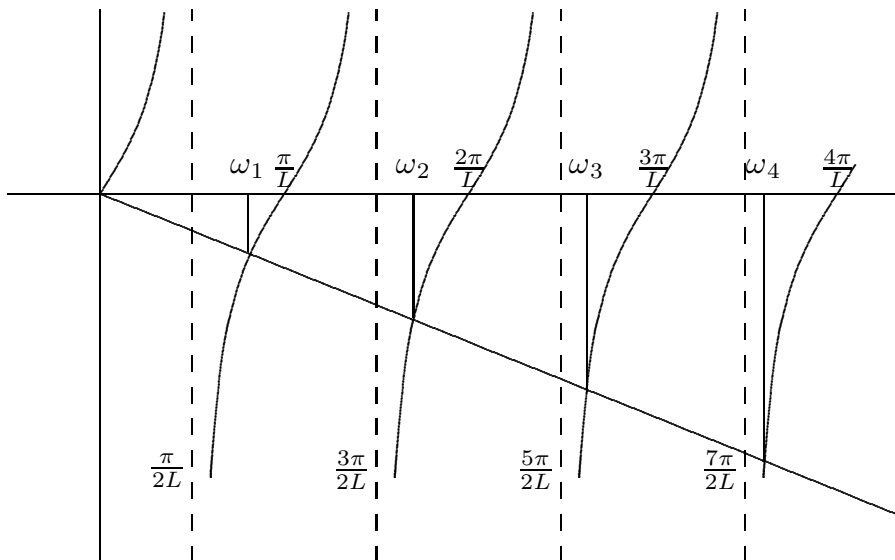
or

$$\frac{\omega}{\beta} \cos \omega L + \sin \omega L = 0, \tag{*}$$

or

$$\tan \omega L = -\frac{1}{\beta} \omega. \tag{*}$$

It is easy to find the *approximate* locations of the eigenvalues,  $\omega^2$ , by graphing the two sides of (\*) (as functions of  $\omega$ ) and picking out the points of intersection. (In the drawing we assume  $\beta > 0$ .)



The  $n$ th root,  $\omega_n$ , is somewhere between  $(n - \frac{1}{2}) \frac{\pi}{L}$  and  $\frac{n\pi}{L}$ ; as  $n \rightarrow \infty$ ,  $\omega_n$  becomes arbitrarily close to  $(n - \frac{1}{2}) \frac{\pi}{L}$ , the vertical asymptote of the tangent function. For smaller  $n$  one could guess  $\omega_n$  by eye and then improve the guess by,

for example, Newton's method. (Because of the violent behavior of  $\tan$  near the asymptotes, Newton's method does not work well when applied to (\*); it is more fruitful to work with (\*)' instead.)

To complete the solution, we write a linear combination of the separated solutions,

$$u(x, t) = \sum_{n=1}^{\infty} b_n \sin \omega_n x e^{-\omega_n^2 t},$$

and seek to determine the coefficients from the initial condition,

$$f(x) = u(x, 0) = \sum_{n=1}^{\infty} b_n \sin \omega_n x.$$

This problem satisfies the conditions of the Sturm–Liouville theorem, so the eigenfunctions

$$\psi_n \equiv \sin \omega_n x$$

are guaranteed to be orthogonal. This can be verified by direct computation (making use of the fact that  $\omega_n$  satisfies (\*)). Thus

$$\int_0^L f(x) \sin \omega_m x \, dx = b_m \int_0^L \sin^2 \omega_m x \, dx.$$

However, the  $\psi_n$  have not been normalized, so we have to calculate

$$\int_0^L \sin^2 \omega_m x \, dx \equiv \|\psi_m\|^2$$

and divide by it. (This number is *not* just  $\frac{1}{2}L$ , as in the Fourier case.) Alternatively, we could construct *orthonormal* basis functions by dividing by the square root of this quantity:

$$\phi_n \equiv \frac{\psi_n}{\|\psi_n\|}.$$

Then the coefficient formula is simply

$$B_m = \int_0^L f(x) \phi_m(x) \, dx$$

(where  $f(x) = \sum_m B_m \phi_m$ , so  $B_m = \|\psi_m\| b_m$ ).

The theorem also guarantees that the eigenfunctions are complete, so this solution is valid for any reasonable  $f$ . (Nevertheless, if  $\beta < 0$  it is easy to overlook one of the normal modes and end up with an incomplete set by mistake. See Haberman, Figs. 5.8.2 and 5.8.3.)



**Continuation of the theorem:** For a regular Sturm–Liouville problem:

- (4) For each eigenvalue  $\omega^2$  there is at most one linearly independent eigenfunction. (Note: This is true only for the “regular” type of boundary conditions,

$$\alpha X'(0) - \beta X(0) = 0, \quad \gamma X'(L) + \delta X(L) = 0.$$

For periodic boundary conditions there can be two independent eigenfunctions for the same  $\omega$ , as we know from Fourier series.)

- (5)  $\omega_n$  approaches  $+\infty$  as  $n \rightarrow \infty$ .
- (6)  $\phi_n(x)$  has exactly  $n - 1$  zeros (“nodes”) in the interval  $(0, L)$  (endpoints not counted). (The basic reason for this is that as  $\omega$  increases,  $\phi$  becomes increasingly concave and oscillatory.)
- (7) If  $\alpha, \beta, \gamma, \delta, p(x)$ , and  $-q(x)$  are all nonnegative, then the  $\omega_n^2$  are all nonnegative. (Corollary: For the heat equation, the solution  $u(x, t)$  approaches 0 as  $t \rightarrow +\infty$  if all the eigenvalues are positive; it approaches a constant if  $\omega = 0$  occurs.)

Note that parts (1) and (7) of the theorem make it possible to exclude the possibilities of complex and negative eigenvalues without a detailed study of the solutions of the ODE for those values of  $\omega^2$ . In first learning about separation of variables and Fourier series we did make such a detailed study, for the ODE  $X'' = -\omega^2 X$ , but I remarked that the conclusion could usually be taken for granted. (Indeed, Appendix A gives the proof of (1) and (7), specialized to  $X'' = -\omega^2 X$ .)

A good exercise: For a regular Sturm–Liouville problem with differential operator

$$\mathcal{L} = \frac{d}{dx} p(x) \frac{d}{dx} + q(x),$$

prove (†) and (7) along the lines previously indicated.

## SINGULAR STURM–LIOUVILLE PROBLEMS

If one of the coefficient functions in the operator  $\mathcal{L}$  violates a condition in the definition of a regular Sturm–Liouville problem at an endpoint (e.g., if  $p(0) = 0$ , or if  $q(x) \rightarrow \infty$  as  $x \rightarrow L$ ), or if the interval is infinite, then the problem is called *singular* (instead of regular). Many of the most important real-life cases are singular. Under these conditions the foregoing theory acquires complications, which I can discuss only very loosely here.

1. The set of eigenfunctions needed to expand an arbitrary function may depend on  $\omega^2$  as a continuous variable, as in the case of the Fourier transform.
2. The boundary conditions needed to get an orthogonal and complete set of eigenfunctions may be of a different type. The critical condition that must be kept satisfied is (†). In particular, if one of the endpoints moves to infinity, then usually there is no boundary condition there of the type  $\gamma X'(L) + \delta X(L) = 0$ ; instead, one merely excludes solutions that grow exponentially fast at infinity. If all the remaining solutions go rapidly to zero at infinity, so that they are square-integrable, then the eigenfunction expansion will be a series, as in the regular problems. If the remaining solutions do not go to zero, then typically *all* of them are needed to form a complete set, and one has a situation like the Fourier transform.

### EIGENFUNCTIONS, DELTA FUNCTIONS, AND GREEN FUNCTIONS

Let's return to the general case and assume that the eigenfunctions have been chosen orthonormal. We have an expansion formula

$$f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x) \quad (\#)$$

and a coefficient formula

$$c_m = \int_a^b \phi_m(z)^* f(z) dz. \quad (b)$$

Substituting (b) into (#) and interchanging the order of summation and integration yields

$$f(x) = \int_a^b dz f(z) \left[ \sum_{n=1}^{\infty} \phi_n(x) \phi_n(z)^* \right].$$

In other words, when acting on functions with domain  $(a, b)$ ,

$$\delta(x - z) = \sum_{n=1}^{\infty} \phi_n(x) \phi_n(z)^*.$$

This is called the *completeness relation* for the eigenfunctions  $\{\phi_n\}$ , since it expresses the fact that the whole function  $f$  can be built up from the pieces  $c_n \phi_n$ . In the special case of the Fourier sine series, we looked at this formula earlier.

We can also substitute (#) into (b), getting

$$c_m = \sum_{n=1}^{\infty} c_n \left[ \int_a^b \phi_m(x)^* \phi_n(x) dx \right].$$

This equation is equivalent to

$$\boxed{\int_a^b \phi_m(x)^* \phi_n(x) dx = \delta_{mn},}$$

where

$$\delta_{mn} \equiv \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n. \end{cases}$$

(This is called the *Kronecker delta symbol*; it is the discrete analogue of the Dirac delta function — or, rather, Dirac's delta function is a continuum generalization of it!) This *orthogonality relation* summarizes the fact that the  $\phi$ 's form an orthonormal basis.

Note that the completeness and orthogonality relations are very similar in structure. Basically, they differ only in that the variables  $x$  and  $n$  interchange roles (along with their alter egos,  $z$  and  $m$ ). The different natures of these variables causes a sum to appear in one case, an integral in the other.

Finally, consider the result of substituting (b) into the solution of an initial-value problem involving the functions  $\phi_n$ . For example, for a certain heat-equation problem we would get

$$u(t, x) = \sum_{n=1}^{\infty} c_n \phi_n(x) e^{-\omega_n^2 t}.$$

This becomes

$$u(t, x) = \int_a^b dz f(z) \left[ \sum_{n=1}^{\infty} \phi_n(x) \phi_n(z)^* e^{-\omega_n^2 t} \right].$$

Therefore, the Green function for that problem is

$$\boxed{G(x, z; t) = \sum_{n=1}^{\infty} \phi_n(x) \phi_n(z)^* e^{-\omega_n^2 t}.$$

When  $t = 0$  this reduces to the completeness relation, since

$$\lim_{t \downarrow 0} G(x, z; t) = \delta(x - z).$$

Similarly,

$$G(x, z; \lambda) = \sum_{n=1}^{\infty} \frac{\phi_n(x) \phi_n(z)^*}{\omega_n^2 - \lambda^2}$$

is the *resolvent kernel*, the Green function such that  $u(x) = \int_0^L G(x, z; \lambda) g(z) dz$  solves the nonhomogeneous ODE  $\mathcal{L}[u] + \lambda^2 u = -g$  (if  $r = 1$ ) with the given boundary

conditions. (Our first Green function example constructed with the aid of the delta function, several sections back, was a resolvent kernel.)

It may be easier to solve for the Green functions directly than to sum the series in these formulas. In fact, such formulas are often used in the reverse direction, to obtain information about the eigenfunctions and eigenvalues from independently obtained information about the Green function.

# Polar Coordinates and Bessel Functions

## POLAR COORDINATES

The polar coordinates  $(r, \theta)$  in  $\mathbf{R}^2$  are defined by

$$\begin{aligned}x &= r \cos \theta, \\y &= r \sin \theta.\end{aligned}$$

The usual reason for rewriting a PDE problem in polar coordinates (or another curvilinear coordinate system) is to make the boundary conditions simpler, so that the method of separation of variables can be applied. For example, the vanishing of  $u(x, y)$  on a circle is easier to apply when expressed as

$$u(4, \theta) = 0$$

than when stated

$$u(x, y) = 0 \quad \text{whenever } x^2 + y^2 = 16.$$

In fact, the latter can't be satisfied by a nontrivial function of the form  $X(x)Y(y)$ , as needed by the separation method.

Indeed, a disc of radius  $r_0$  is, in polar coordinates, the region

• disc: 
$$0 \leq r < r_0, \quad 0 \leq \theta < 2\pi.$$

It is the most obvious of the types of regions that “look like rectangles” when expressed in polar coordinates. Others are

• exterior of disc: 
$$0 < r_0 < r < \infty, \quad 0 \leq \theta < 2\pi;$$

• annulus: 
$$0 < r_1 < r < r_2, \quad 0 \leq \theta < 2\pi;$$

• sector: 
$$0 \leq r < r_0, \quad \Theta_1 \leq \theta < \Theta_2;$$

and three others that have no convenient names (although “partially eaten piece of pie” might do for one of them).

In any such case one will want to rewrite the whole problem in polar coordinates to exploit the geometry. This is likely to make the PDE itself more complicated, however. At least once in your life, you should go through the calculation — using the product rule and multivariable chain rule repeatedly, starting from formulas such as

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x} \frac{\partial}{\partial \theta}$$

— that shows that the two-dimensional Laplacian operator

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

is equal to

$$\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}.$$

It is worth noting that the  $r$ -derivative terms

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r}$$

can also be written as a single term,

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right).$$

#### SEPARATION OF VARIABLES IN THE POLAR POTENTIAL EQUATION

Let us, therefore, study Laplace's equation

$$0 = \nabla^2 u = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

We try separation of variables:

$$u(r, \theta) = R(r)\Theta(\theta).$$

We get

$$\frac{1}{r}(rR')'\Theta + \frac{1}{r^2}R\Theta'' = 0$$

(where the primes are unambiguous, because each function depends on only one variable). Observe that we can separate the  $r$  and  $\theta$  dependence into different terms by dividing by  $R\Theta/r^2$ :

$$\frac{r(rR)'}{R} + \frac{\Theta''}{\Theta} = 0.$$

We can therefore introduce an unknown constant (eigenvalue) and split the equation into two ordinary DEs:

$$\frac{\Theta''}{\Theta} = K, \quad \frac{r(rR)'}{R} = -K.$$

The first of these is our old friend whose solutions are the trig functions; we put it aside to deal with later.

More interesting is the radial equation,

$$(rR')' + \frac{K}{r} R = 0$$

or

$$R'' + \frac{1}{r} R' + \frac{K}{r^2} R = 0.$$

It is of the general Sturm–Liouville type. Consulting the theorems and definitions concerning those, we see that we will have a *regular* Sturm–Liouville problem provided that the boundaries of our region are

$$r = r_1 \quad \text{and} \quad r = r_2 \quad \text{with } r_1 > 0 \text{ and } r_2 < \infty$$

— that is, for the half-eaten piece of pie and the annulus (ring). For the more common situations of the disc, disc exterior, and sector, the SL problem is *singular*.

However, a little learning is a dangerous thing. Although the analysis I have just given you is correct, and will be valuable soon when we complicate the equation by adding another term, it turns out to be unnecessary in the present case. Let's make the change of variables

$$z \equiv \ln r \quad (\text{hence } r = e^z),$$

so that

$$\frac{d}{dr} = \frac{dz}{dr} \frac{d}{dz} = \frac{1}{r} \frac{d}{dz}.$$

Then

$$rR' = \frac{R}{z}, \quad (rR')' = \frac{1}{r} \frac{d^2 R}{dz^2},$$

so the equation becomes

$$\frac{d^2 R}{dz^2} + KR = 0.$$

It is our old friend after all!

Let us record its (basis) solutions for the various classes of  $K$ :

1.  $K = -\lambda^2 < 0$ :  $R = e^{\pm\lambda z} = r^{\pm\lambda}$ .
2.  $K = 0$ :  $R = 1$  and  $R = z = \ln r$ .
3.  $K = \mu^2 > 0$ :  $R = e^{\pm i\mu z} = r^{\pm i\mu}$ ;

that is,  $R = \cos(\mu \ln r)$  and  $R = \sin(\mu \ln r)$ .

We consider two examples and make brief remarks on a third.

**I. Interior of a disc of radius  $r_0$**

Three boundary conditions, of different natures, arise here.

First, since the coordinate  $\theta$  goes “all the way around”,  $u(r, \theta)$  must be periodic in  $\theta$  with period  $2\pi$ . Therefore, the solutions of the angular equation,  $\Theta'' = K\Theta$ , will be the terms of a full Fourier series at the standard scale:

$$u(r, \theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} R_n(r).$$

(Of course, we could use sines and cosines instead.) Moreover,  $K = -n^2$ .

Second, at the rim of the disc a well-posed potential problem requires a standard nonhomogeneous Dirichlet, Neumann, or Robin condition, such as

$$u(r_0, \theta) = f(\theta).$$

This will be applied to the whole series, not each term  $R_n$ , and will eventually determine the coefficients  $c_n$ .

Third, to complete the specification of  $R_n$  we need to say how the solution behaves as  $r \rightarrow 0$ . We know that  $u(r, \theta)$  reexpressed as a function of  $x$  and  $y$  must be a *solution* at the center of the disc. This implies that  $R_n(r)$  must stay bounded as  $r$  approaches 0. Looking back at our list of possible radial solutions, we see that the allowable ones are  $R_0(r) = 1$  and  $R_n(r) = r^{|n|}$  for  $n \neq 0$ .

So, finally, the solution is

$$u(r, \theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} r^{|n|},$$

where  $c_n$  must be determined by (in the Dirichlet case)

$$f(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} r_0^{|n|};$$

that is,

$$c_n = \frac{1}{r_0^{|n|}} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta) d\theta.$$



## II. A partially eaten piece of pie

Consider the truncated sector, or polar rectangle, bounded by the four curves

$$r = r_1, \quad r = r_2, \quad \theta = \theta_1, \quad \theta = \theta_2,$$

where  $0 < r_1$  and  $r_2 < \infty$ . In this case, all four boundaries are of the “regular” type. Let’s suppose that nonhomogeneous data are given on all four sides — something like

$$\begin{aligned} u(r_1, \theta) &= f_1(\theta), & u(r_2, \theta) &= f_2(\theta), \\ \frac{\partial u}{\partial \theta}(r, \theta_1) &= f_3(r), & \frac{\partial u}{\partial \theta}(r, \theta_2) &= f_4(r). \end{aligned}$$

As in the Cartesian rectangle case, before separating variables we must split this into two problems, one with homogeneous  $\theta$  boundary conditions and one with homogeneous  $r$  boundary conditions. Let us say  $u = v + w$ , where  $v$  and  $w$  individually solve the potential equation,  $v$  satisfies

$$v(r_1, \theta) = 0, \quad v(r_2, \theta) = 0, \quad \frac{\partial v}{\partial \theta}(r, \theta_1) = f_3(r), \quad \frac{\partial v}{\partial \theta}(r, \theta_2) = f_4(r),$$

and  $w$  satisfies

$$w(r_1, \theta) = f_1(\theta), \quad w(r_2, \theta) = f_2(\theta), \quad \frac{\partial w}{\partial \theta}(r, \theta_1) = 0, \quad \frac{\partial w}{\partial \theta}(r, \theta_2) = 0.$$

In solving for  $v$ , it is the homogeneous conditions on  $R$  that must determine the allowed eigenvalues  $K$ . Thus here, for the first time, we really treat the radial equation as a Sturm–Liouville problem. In order for  $R$  to vanish at both  $r_1$  and  $r_2$ , we must have  $K > 0$ , the third case in our list of radial solutions. That is, for each normal mode we have an eigenvalue  $K_\mu = \mu^2$  and an eigenfunction

$$R_\mu(r) = A_\mu \cos(\mu \ln r) + B_\mu \sin(\mu \ln r)$$

(or, alternatively,  $R_\mu = C_{\mu,+} r^{i\mu} + C_{\mu,-} r^{-i\mu}$ ). The two equations

$$R(r_1) = 0 = R(r_2)$$

(1) determine a discrete list of allowable values of  $\mu$ , and (2) determine the ratio of  $A_\mu$  to  $B_\mu$  (or  $C_{\mu,+}$  to  $C_{\mu,-}$ ). This leaves an overall constant factor in  $R_\mu$  undetermined, as is always the case in finding normal modes. I postpone the details of this calculation for a moment; the principle is the same as in the very first separation-of-variables problem we did, where the eigenvalues turned out to be  $(n\pi/L)^2$  and the eigenfunctions  $\sin(n\pi x/L)$  times an arbitrary constant.

To finish the solution for  $v$  we need to solve the equation

$$\Theta'' = +\mu^2 \Theta.$$

Thus the angular dependence of this solution is exponential, not trigonometric. We can write

$$v(r, \theta) = \sum_{\mu} c_{\mu} R_{\mu}(r) (C_{\mu} e^{\mu\theta} + D_{\mu} e^{-\mu\theta}).$$

The constants  $C$  and  $D$  are to be determined by imposing the remaining boundary conditions,

$$\frac{\partial v}{\partial \theta}(r, \theta_1) = f_3(r), \quad \frac{\partial v}{\partial \theta}(r, \theta_2) = f_4(r).$$

In general this will be a coupled pair of Sturm–Liouville expansions in the orthogonal eigenfunctions  $R_{\mu}(r)$ .

That's  $v$ ; now we need to find  $w$ . That problem is like this one, except that the roles of  $r$  and  $\theta$  are interchanged. The result will be a Fourier cosine series in  $\theta$  with radial factors that depend exponentially on  $\ln r$ ; that is, linear combinations of  $r^n$  and  $r^{-n}$ . I hope that by now I can leave the details to your imagination.

*Unfinished business:* Let us consider the details of finding the eigenvalues  $\mu$  and eigenfunctions  $R_{\mu}$ . The two relevant algebraic equations are

$$0 = R_{\mu}(r_1) = C_{\mu,+} r_1^{i\mu} + C_{\mu,-} r_1^{-i\mu}$$

and

$$0 = R_{\mu}(r_2) = C_{\mu,+} r_2^{i\mu} + C_{\mu,-} r_2^{-i\mu}.$$

A nontrivial solution will exist if and only if the determinant vanishes:

$$0 = \begin{vmatrix} r_1^{i\mu} & r_1^{-i\mu} \\ r_2^{i\mu} & r_2^{-i\mu} \end{vmatrix} = \left(\frac{r_1}{r_2}\right)^{i\mu} - \left(\frac{r_1}{r_2}\right)^{-i\mu}.$$

This is proportional to

$$\sin(\mu \ln(r_1/r_2)),$$

so it vanishes precisely when  $\mu$  is an integer multiple of the constant  $\pi / \ln(r_1/r_2)$ .

Returning to one of the linear algebraic equations, we find

$$\frac{C_{\mu,+}}{C_{\mu,-}} = -r_1^{-2i\mu} = -e^{-2i\mu \ln r_1}.$$

(Using the other equation would give  $C_{\mu,+}/C_{\mu,-} = -e^{-2i\mu \ln r_2}$ , but these two equations are equivalent because of the eigenvalue condition, which may be rewritten as  $\mu \ln r_1 - \mu \ln r_2 = N\pi$ .) The neatest (albeit not the most obvious) normalization convention is to choose  $C_{\mu,-} = -r_1^{i\mu}$ ; then  $C_{\mu,+} = r_1^{-i\mu}$ , and

$$\begin{aligned} R_{\mu}(r) &\equiv C_{\mu,+} r^{i\mu} + C_{\mu,-} r^{-i\mu} \\ &= \left(\frac{r}{r_1}\right)^{i\mu} - \left(\frac{r}{r_1}\right)^{-i\mu} \\ &= 2i \sin(\mu \ln(r/r_1)). \end{aligned}$$

Thus the Sturm–Liouville expansion involved in this problem is simply an ordinary Fourier sine series, though expressed in very awkward notation because of the context in which it arose. (In our usual notation, we have  $L \equiv \ln(r_1/r_2)$ ,  $\mu = n\pi/L$ ,  $x = z - \ln r_1$ ,  $C_{\mu,+} = b_n/2i$ .) We would have encountered the same complications in Cartesian coordinates if we had considered examples where none of the boundaries lay on the coordinate axes (but the boundaries were parallel to the axes).

### III. A sector (the pie intact)

Consider the region

$$0 \leq r < r_2, \quad \Theta_1 \leq \theta < \Theta_2,$$

with nonhomogeneous data on the straight sides. (This is the limiting case of the  $v$  problem above as  $r_1 \rightarrow 0$ .) The endpoint  $r = 0$  is singular, so we are not guaranteed that a standard Sturm–Liouville expansion will apply. Indeed, in terms of the variable  $z = \ln r$ , where the radial equation becomes trivial, the endpoint is at  $z = -\infty$ . This problem is therefore a precise polar analogue of the infinite rectangular slot problem, and the solution will be a Fourier sine or cosine *transform* in a variable  $\zeta \equiv -z + C$  that vanishes when  $r = r_2$ . (That is,  $C = \ln r_2$ .)

## BESSEL FUNCTIONS

THE DRUM PROBLEM: Consider the wave equation (with  $c = 1$ ) in a disc with homogeneous Dirichlet boundary conditions:

$$\begin{aligned} \nabla^2 u &= \frac{\partial^2 u}{\partial t^2}, & u(r_0, \theta, t) &= 0, \\ u(r, \theta, 0) &= f(r, \theta), & \frac{\partial u}{\partial t}(r, \theta, 0) &= g(r, \theta). \end{aligned}$$

(Note that to solve the *nonhomogeneous* Dirichlet problem for the wave equation, we would add this solution to that of the disc potential problem, **I**, solved in the previous section; the latter is the steady-state solution for the wave problem.)

We expect to get a sum over normal modes,

$$u = \sum_n \phi_n(r, \theta) T_n(t).$$

Let us seek the separated solutions: If  $u_{\text{sep}} = \phi(r, \theta)T(t)$ , then

$$\frac{\nabla^2 \phi}{\phi} = \frac{T''}{T} = -\omega^2.$$

Therefore

$$T = \alpha e^{i\omega t} + \beta e^{-i\omega t} = A \cos(\omega t) + B \sin(\omega t).$$

As for  $\phi$ , it will be periodic in  $\theta$  and satisfy  $\phi(r_0, \theta) = 0$  along with the *Helmholtz equation*

$$-\omega^2 \phi = \nabla^2 \phi = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2}.$$

This is still a partial DE, so we separate variables again:  $\phi = R(r)\Theta(\theta)$ ,

$$\frac{r(rR)'}{R} + r^2\omega^2 = -\frac{\Theta''}{\Theta} = \nu^2.$$

(In the last section we did this step for  $\omega = 0$ , and  $\nu^2$  was called  $-K$ .) The boundary condition becomes  $R(r_0) = 0$ , and as in the previous disc problem we need to assume that  $R$  is bounded as  $r \rightarrow 0$ , so that  $\phi$  will be differentiable at the origin and be a solution there. The angular equation is the familiar  $\Theta'' = -\nu^2\Theta$ , with solutions

$$\Theta(\theta) = e^{\pm in\theta} \quad \text{with } n = |\nu| \text{ an integer.}$$

*Remark:* Unlike the last disc problem, here we have homogeneous BC on *both*  $\Theta$  and  $R$ . The nonhomogeneity in this problem is the initial data on  $u$ .

We can write the radial equation in the Sturm–Liouville form

$$(rR')' - \frac{n^2}{r} R + \omega^2 r R = 0$$

or in the form

$$R'' + \frac{1}{r} R' + \left( \omega^2 - \frac{\nu^2}{r^2} \right) R = 0.$$

This is called *Bessel's equation* if  $\omega^2 \neq 0$ . (We already studied the case  $\omega = 0$  at length. Recall that the solutions were powers of  $r$ , except that  $\ln r$  also appeared if  $n = 0$ .) We can put the Bessel equation into a standard form by letting

$$z \equiv \omega r; \quad r = \frac{z}{\omega}, \quad \frac{d}{dr} = \omega \frac{d}{dz}.$$

After dividing by  $\omega^2$  we get

$$\boxed{\frac{d^2 R}{dz^2} + \frac{1}{z} \frac{dR}{dz} + \left( 1 - \frac{n^2}{z^2} \right) R = 0.}$$

(The point of this variable change is to get an equation involving only one arbitrary parameter instead of two.)

If we have a solution of this equation, say  $R = Z_n(z)$ , then  $R(r) \equiv Z_n(\omega r)$  is a solution of the original equation (with  $\nu = \pm n$ ). All solutions  $Z_n(z)$  are called *Bessel functions of order  $n$* . Although they are not expressible in terms of elementary functions (except when  $n$  is half an odd integer), they have been studied so much that many properties of them are known and tabulated in handbooks, symbolic algebra programs, etc.

*Remark:* For the disk problem,  $n$  must be an integer (which we can take nonnegative), but for *sector* problems, other values of  $n$  can appear.

## PROPERTIES OF BESSEL FUNCTIONS

### Power series solution:

In a good differential equations course one learns to substitute

$$Z_n(z) = z^\alpha \sum_{m=0}^{\infty} c_m z^m$$

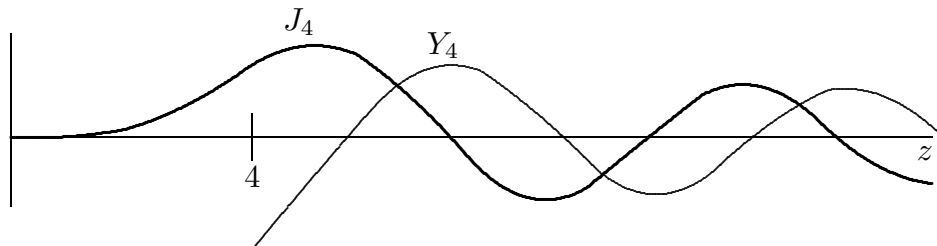
into Bessel's equation, equate the coefficient of each power of  $z$  to 0, and try to solve for  $\alpha$  and the  $c_m$  ("method of Frobenius"). It turns out that  $\alpha = \pm n$  (so  $\alpha$  can be identified with the  $\nu$  of the original equation), and that for  $n$  a nonnegative integer there is a solution of the assumed form only for the positive root. It is called " $J$ ":

$$J_n(z) \equiv \left(\frac{z}{2}\right)^n \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(n+m)!} \left(\frac{z}{2}\right)^{2m}.$$

This series converges for all  $z$ .

Any solution linearly independent of  $J_n$  has a singularity at  $z = 0$  (in fact, it goes to  $\infty$  in absolute value there). For noninteger  $n$  the series with  $\alpha = -n$  exists and contains negative powers, but for integer  $n$  the second solution involves a logarithm. (It can be found by the method of "reduction of order".) This second solution is nonunique, because of the freedom to multiply by a constant and the freedom to add a multiple of  $J_n$ . However, there is a standard choice ("normalization") of the second solution, called either  $Y_n(z)$  or  $N_n(z)$ ; I prefer " $Y$ ".

**General behavior:** Here is a graph of  $J_4$  and  $Y_4$ . Near the origin,  $J_n$  behaves like  $z^n$ , while  $Y_n$  blows up like  $z^{-n}$  (like  $\ln z$  if  $n = 0$ ). At large  $z$  both functions oscillate, with a slowly decreasing amplitude.



**Behavior at small argument ( $z \rightarrow 0$ ):**

Think of  $J_n$  as like  $r^n$ ,  $Y_n$  as like  $r^{-n}$ . More precisely,

$$J_n(z) \approx \frac{1}{n!} \left(\frac{z}{2}\right)^n,$$

$$Y_0(z) \approx \frac{2}{\pi} \ln z,$$

$$Y_n(z) \approx -\frac{(n-1)!}{\pi} \left(\frac{z}{2}\right)^{-n} \quad \text{if } n > 0.$$

Therefore, for a problem inside a disc only  $J$  functions will appear, by the boundedness criterion previously mentioned.

**Behavior at large argument ( $z \rightarrow +\infty$ ):**

Think of  $J_n$  as like  $\cos$ ,  $Y_n$  as like  $\sin$ . More precisely,

$$J_n(z) \approx \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{1}{2}n\pi - \frac{1}{4}\pi\right),$$

$$Y_n(z) \approx \sqrt{\frac{2}{\pi z}} \sin\left(z - \frac{1}{2}n\pi - \frac{1}{4}\pi\right).$$

One defines the analogues of complex exponentials:

$$H_n^{(1)}(z) \equiv J_n + iY_n \approx \sqrt{\frac{2}{\pi z}} (-i)^{n+\frac{1}{2}} e^{iz},$$

$$H_n^{(2)}(z) \equiv J_n - iY_n \approx \sqrt{\frac{2}{\pi z}} i^{n+\frac{1}{2}} e^{-iz}.$$

The crossover point between the  $r^{\pm n}$  behavior and the trigonometric behavior is somewhere close to  $z = n$ .

It is *not* necessary to memorize all these formulas. You should know:

1.  $J$  is bounded and smooth at 0;  $Y$  isn't.
2. The Bessel functions (for real  $n$  and  $\omega$ ) are oscillatory at infinity. (Note that their "envelope" decreases as  $1/\sqrt{z}$ , but this is not enough to make them square-integrable.)

**Recursion relations:**

$$\begin{aligned}zJ'_n + nJ_n &= zJ_{n-1}, \\zJ'_n - nJ_n &= -zJ_{n+1}.\end{aligned}$$

From these follow

$$\frac{2n}{z} J_n = J_{n-1} + J_{n+1}$$

and, most useful of all,

$$2J'_n = J_{n-1} - J_{n+1}.$$

(So the *derivative* of a Bessel function is not really a new function. Note that the *second* (and hence any higher) derivative can be calculated using the Bessel equation itself.)

The recursion relations are useful in many ways. For instance, computer programs need to calculate  $J_n$  “by brute force” only for a few values of  $n$  and then use the recursion relations to interpolate.

MODIFIED BESSEL FUNCTIONS (AND OTHER SUCH THINGS)

In the application just discussed, we had  $\nu^2 > 0$  and  $\omega^2 > 0$ . But Bessel’s equation,

$$R'' + \frac{1}{r} R' + \left( \omega^2 - \frac{\nu^2}{r^2} \right) R = 0.$$

also makes sense, and has applications, when one or both of these parameters is negative or complex, so that  $\nu$  or  $\omega$  is complex. Complex  $\omega$  corresponds to complex  $z$ , since  $z = \omega r$ . In particular, imaginary  $\omega$  (negative real  $\omega^2$ ) corresponds to evaluation of the Bessel functions on the imaginary axis:  $Z_\nu(i|\omega|r)$ . This is analogous to the passage from  $e^{\pm nx}$  to  $e^{\pm inx}$ , which yields the trigonometric functions (except that here we are moving in the reverse direction, as we shall now see).

These Bessel functions of imaginary argument (but real  $\nu$ ) are called *modified Bessel functions*. A standard basis consists of two functions called  $I_\nu(z)$  and  $K_\nu(z)$ , chosen to behave somewhat like  $\sinh z$  and  $e^{-z}$ , respectively.

**Definitions:**

$$\begin{aligned}I_\nu(z) &\equiv i^{-\nu} J_\nu(iz), \\K_\nu(z) &\equiv \frac{\pi}{2} i^{\nu+1} H_\nu^{(1)}(iz).\end{aligned}$$

**Behavior at small argument ( $z \rightarrow 0$ ):**

$$\begin{aligned} I_\nu(z) &\approx \frac{1}{\nu!} \left(\frac{z}{2}\right)^\nu, \\ K_\nu(z) &\approx \frac{1}{2} (\nu - 1)! \left(\frac{z}{2}\right)^{-\nu}, \\ K_0(z) &\approx -\ln z. \end{aligned}$$

**Behavior at large argument ( $z \rightarrow +\infty$ ):**

$$\begin{aligned} I_\nu(z) &\approx \frac{e^z}{\sqrt{2\pi z}}, \\ K_\nu(z) &\approx \sqrt{\frac{\pi}{2z}} e^{-z}. \end{aligned}$$

In summary,  $I_\nu$  is designed to vanish at 0, whereas  $K_\nu$  is designed to vanish at infinity. (But the arbitrary constant factors in the definitions arose by historical accidents that are not worth wondering about.)

An application of modified Bessel functions will be given later.

Bessel functions of imaginary order  $\nu$  appear in separation of variables in *hyperbolic coordinates*,

$$\begin{array}{ccc} t = r \sinh \theta, & & t = r \cosh \theta, \\ x = r \cosh \theta & \text{or} & x = r \sinh \theta. \end{array}$$

(The first of these transformations of variables can be related to the “twin paradox” in special relativity. The two apply to different regions of the  $t$ - $x$  plane.) If you apply such a transformation to the Klein–Gordon equation,

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} + m^2 u = 0,$$

you will get for the  $r$  dependence a Bessel equation with imaginary  $\nu$  and real or imaginary  $\omega$  (depending on which of the two hyperbolic transformations you’re using). Therefore, the solutions will be either  $J_{i\kappa}$  or  $K_{i\kappa}$  functions.

Many ordinary differential equations are Bessel’s equation in disguise. That is, they become Bessel’s equation after a change of dependent or independent variable, or both. One example is the deceptively simple-looking equation

$$\frac{d^2 u}{dx^2} + xu = 0,$$



whose solutions are called *Airy functions*. If you let

$$y \equiv \frac{2}{3}x^{\frac{3}{2}}, \quad u \equiv x^{\frac{1}{2}}Z,$$

then you get

$$\frac{d^2Z}{dy^2} + \frac{1}{y} \frac{dZ}{dy} + \left(1 - \frac{1}{9y^2}\right)Z = 0,$$

the Bessel equation of order  $\nu = \frac{1}{3}$ . Therefore, the Airy functions are essentially Bessel functions:

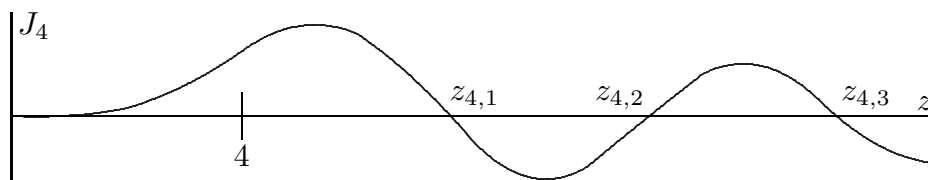
$$u = \sqrt{x} Z_{\frac{1}{3}}\left(\frac{2}{3}x^{\frac{3}{2}}\right).$$

### FINISHING UP THE DRUM PROBLEM

Recall that we were seeking normal modes  $\phi_\nu(r, \theta) = R(r)\Theta(\theta)$ , where

$$\Theta(\theta) = e^{i\nu\theta} \equiv e^{\pm i n \theta} \quad \text{with } \nu \text{ an integer}$$

and  $R(r) = Z_n(\omega r)$  had to be a Bessel function satisfying appropriate boundary conditions at the origin and the edge of the disk ( $r = r_0$ ). As we have seen, the condition that  $R$  remain bounded as  $r \rightarrow 0$  implies that  $Z_n$  is (a multiple of)  $J_n$ . The other condition is that  $R(r_0) = 0$ . It follows that  $\omega r_0$  must equal a zero (root) of  $J_n(z)$ . Since  $J_n$  eventually becomes oscillatory, there are infinitely many such values of  $z$ ; let us call them  $z_{n1}, z_{n2}, \dots$ . They can be found by numerical methods and are tabulated in many handbooks. They are the analogue for the current problem of the numbers  $n\pi/L$  in simple Fourier problems and the roots of the equations such as  $\tan z = -\gamma z$  in convective Sturm–Liouville problems.



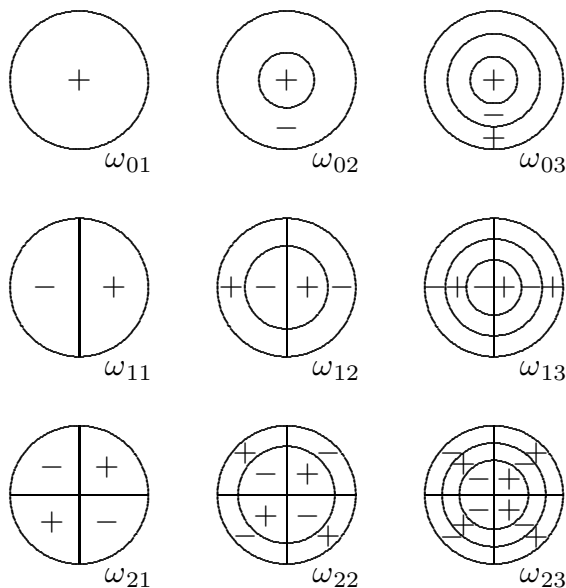
Therefore, the eigenvalues of our problem (or, rather, their square roots) are

$$\omega_{nk} \equiv \frac{z_{nk}}{r_0}. \quad (1)$$

The presence of  $\omega_{nk}$  scaling the radial coordinate “compresses” the  $n$ th Bessel function so that  $k$  of the lobes of its graph fit inside the disk of radius  $r_0$ . Putting the radial and angular parts together, we have the eigenfunctions

$$\psi_{\nu k}(r, \theta) = R(r)\Theta(\theta) = J_n(\omega_{nk}r) e^{i\nu\theta} \quad (\nu = \pm n). \quad (2)$$

We could equally well use the *real* eigenfunctions in which  $e^{i\nu\theta}$  is replaced by  $\sin n\theta$  or  $\cos n\theta$ ; those functions are easier to visualize. In the drawing the lines and curves indicate places where such a  $\psi$  equals 0, and the signs indicate how the solution  $\text{Re } \psi$  or  $\text{Im } \psi$  bulges above or below the plane  $\psi = 0$ . Such patterns may be seen in the surface of a cupful of coffee or other liquid when the container is tapped lightly. (Compare the rectangle eigenfunctions in an earlier section.)



If we were solving a heat-conduction problem in the disk, the general solution would be a linear combination of the separated solutions:

$$u(t, r, \theta) = \sum_{\nu=-\infty}^{\infty} \sum_{k=1}^{\infty} c_{\nu k} \psi_{\nu k}(r, \theta) e^{-\omega_{nk}^2 t}. \quad (3)$$

The coefficients need to be calculated from the initial data:

$$\begin{aligned} g(r, \theta) &= u(0, r, \theta) \\ &= \sum_{\nu=-\infty}^{\infty} \sum_{k=1}^{\infty} c_{\nu k} \psi_{\nu k}(r, \theta) \\ &= \sum_{\nu=-\infty}^{\infty} \sum_{k=1}^{\infty} c_{\nu k} J_n(\omega_{nk} r) e^{i\nu\theta}. \end{aligned}$$

By the standard Fourier series formula,

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-i\nu\theta} g(r, \theta) d\theta = \sum_{k=1}^{\infty} c_{\nu k} J_n(\omega_{nk} r).$$

We are left with a one-dimensional series in the eigenfunctions  $R_n(r) \equiv J_n(\omega_{nk} r)$  ( $n$  fixed). We recall that these functions came out of the equation

$$R'' + \frac{1}{r} R' + \left( \omega^2 - \frac{n^2}{r^2} \right) R = 0$$

with the boundary condition  $R(r_0) = 0$ , which looks like a Sturm–Liouville problem. Unfortunately, it does not quite satisfy the technical conditions of the Sturm–Liouville theorem, because of the singular point in the ODE at  $r = 0$ . Nevertheless,

it turns out that the conclusions of the theorem are still valid in this case: The eigenfunctions are complete (for each fixed  $n$ ), and they are orthogonal with respect to the weight function  $r$ :

$$\int_0^{r_0} J_n(\omega_{ni}r) J_n(\omega_{nj}r) r dr = 0 \quad \text{if } i \neq j.$$

Thus if  $h(r)$  is an arbitrary function on  $[0, r_0]$ , it can be expanded as

$$h(r) = \sum_{k=1}^{\infty} c_k J_n(\omega_{nk}r),$$

and the coefficients are

$$c_k = \frac{\int_0^{r_0} J_n(\omega_{nk}r) h(r) r dr}{\int_0^{r_0} J_n(\omega_{nk}r)^2 r dr}.$$

Furthermore, the experts on Bessel functions assure us that the integral in the denominator can be evaluated:

$$\int_0^1 J_n(z_{nk}\zeta)^2 \zeta d\zeta = \frac{1}{2} J_{n+1}(z_{nk})^2.$$

(I leave the change of variable from  $\zeta$  to  $r$  as an exercise.)

Applying this theorem to our problem, we get

$$c_{\nu k} = \left[ \frac{r_0^2}{2} J_{n+1}(\omega_{nk})^2 \right]^{-1} \int_0^{r_0} r dr J_n(\omega_{nk}r) \frac{1}{2\pi} \int_0^{2\pi} d\theta e^{-i\nu\theta} g(r, \theta). \quad (4)$$

That is,

$$\begin{aligned} c_{\nu k} &= \left[ \pi r_0^2 J_{n+1}(\omega_{nk})^2 \right]^{-1} \int_{r=0}^{r_0} \int_{\theta=0}^{2\pi} r dr d\theta \psi_{\nu k}(r, \theta)^* g(r, \theta) \\ &= \frac{1}{\|\psi_{\nu k}\|^2} \int_{r=0}^{r_0} \int_{\theta=0}^{2\pi} r dr d\theta \psi_{\nu k}(r, \theta)^* g(r, \theta). \end{aligned}$$

(In the last version I have identified the constant factor as the normalization constant for the *two-dimensional* eigenfunction.) We now see that the mysterious weight factor  $r$  has a natural geometrical interpretation: It makes the  $r$  and  $\theta$  integrations go together to make up the standard integration over the disc in polar coordinates!

The formulas (1)–(4) give a complete solution of the heat-conduction problem.

But I thought we were solving the wave equation, to model the vibrations of a drum? Yes, your absent-minded professor shifted to the heat equation in midstream, then decided to stay there to keep the formulas simpler. What changes are needed in the foregoing to finish the drum problem? The eigenvalues (1) and eigenfunctions (2) are the same. However, for each eigenfunction there are now *two* possible terms in the solution; the eigenfunction expansion (3) needs to be replaced by

$$u(t, r, \theta) = \sum_{\nu=-\infty}^{\infty} \sum_{k=1}^{\infty} c_{\nu k} \psi_{\nu k}(r, \theta) \cos(\omega_{nk}t) + d_{\nu k} \psi_{\nu k}(r, \theta) \sin(\omega_{nk}t).$$

[There is an important pitfall to avoid here, which is not confined to polar coordinates. (It also arises, for instance, in the wave equation for vibrations in a ring, using Fourier series.) Suppose that you chose to use the real eigenfunctions. Then it would be a mistake to write in the summand something like

$$[a_{nk} \cos(n\theta) + b_{nk} \sin(n\theta)][c_{nk} \cos(\omega_{nk}t) + d_{nk} \sin(\omega_{nk}t)].$$

This would result in equations for the unknown coefficients that are nonlinear, hence hard to solve; also, the solution will not be unique, and may not even exist for some initial data. Remember to write the general solution as a *linear combination* of all possible (independent) elementary separated solutions:

$$A_{nk} \cos(n\theta) \cos(\omega_{nk}t) + B_{nk} \cos(n\theta) \sin(\omega_{nk}t) \\ + C_{nk} \sin(n\theta) \cos(\omega_{nk}t) + D_{nk} \sin(n\theta) \sin(\omega_{nk}t).$$

In other words, multiply first, *then* superpose!]

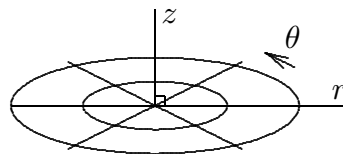
To finish the problem, we need to set  $u$  and its time derivative equal to the given initial data and solve for the  $c$  and  $d$  coefficients. The same orthogonality properties used in the treatment of the heat equation apply here, so (after twice as much work) you will end up with formulas analogous to (4).

#### A HIGHER-DIMENSIONAL EXAMPLE

We shall consider the three-dimensional potential equation in a cylinder. (See J. D. Jackson, *Classical Electrodynamics*, Chapter 3.)

Cylindrical coordinates are defined by

$$x = r \cos \theta, \\ y = r \sin \theta, \\ z = z.$$

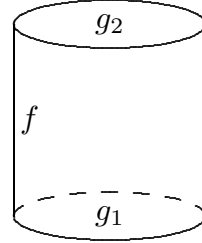


The Laplacian operator is

$$\begin{aligned}\nabla^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\ &= \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2}.\end{aligned}$$

The problem to solve is:  $\nabla^2 u = 0$  inside the cylinder, with Dirichlet data given on all three parts of the cylinder's surface:

$$\begin{aligned}u(r_0, \theta, z) &= f(\theta, z), \\ u(r, \theta, 0) &= g_1(r, \theta), \\ u(r, \theta, L) &= g_2(r, \theta).\end{aligned}$$



As usual, the first step is to split this into two problems:

$$u = u_1 + u_2,$$

where

$$u_1(r_0, \theta, z) = 0 \quad \text{with nonhomogeneous data on the end faces,}$$

$$u_2(r, \theta, 0) = 0 = u_2(r, \theta, L) \quad \text{with nonhomogeneous data on the curved surface.}$$

In either of these subproblems we can separate variables this way:

$$u = R(r)\Theta(\theta)Z(z).$$

After several routine steps (exercise) we get

$$\frac{d^2 Z}{dz^2} - \omega^2 Z = 0, \quad \frac{d^2 \Theta}{d\theta^2} + \mu^2 \Theta = 0,$$

$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left( \omega^2 - \frac{\mu^2}{r^2} \right) R = 0$$

except that it is not yet clear whether the quantities here named  $\omega^2$  and  $\mu^2$  are really positive. (If we find out they aren't, we'll change notation.) Note that the radial equation is a Bessel equation.

*Problem 1:* In the  $u_1$  problem the homogeneous boundary condition is  $R(r_0) = 0$ . The equations determining  $\Theta$  and  $R$  are identical to these we solved in the drum problem. So, we have  $\mu = \pm n$ , an integer, and then  $\omega_{nk} = z_{nk}/r_0$ , where  $z_{nk}$

is the  $k$ th zero of  $J_n$ . The new element is the  $Z$  equation, whose solutions are exponentials. As in some previous problems, the most convenient basis for these solutions consists of certain hyperbolic functions. Cutting a long story short, we arrive at the general solution

$$u_1(r, \theta, z) = \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} J_n(\omega_{nk}r) [A_{nk} \cos n\theta \sinh \omega_{nk}z + B_{nk} \sin n\theta \sinh \omega_{nk}z \\ + C_{nk} \cos n\theta \sinh \omega_{nk}(L-z) + D_{nk} \sin n\theta \sinh \omega_{nk}(L-z)].$$

(I chose real eigenfunctions for variety, and to reinforce an earlier warning about how to write correct linear combinations of normal modes.) Then, for example, we have

$$g_1(r, \theta) = u_1(r, \theta, 0) \\ = \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} J_n(\omega_{nk}r) [C_{nk} \cos(n\theta) \sinh(\omega_{nk}L) + D_{nk} \sin(n\theta) \sinh(\omega_{nk}L)],$$

and therefore

$$C_{nk} = \frac{\int_0^{r_0} r dr \int_0^{2\pi} d\theta J_n(\omega_{nk}r) \cos(n\theta) g_1(r, \theta)}{\pi \sinh(\omega_{nk}L) \int_0^{r_0} J_n(\omega_{nk}r)^2 r dr}.$$

The solutions for  $C_{0k}$  and  $D_{nk}$ , and the solutions for  $A_{nk}$  and  $B_{nk}$  in terms of  $g_2$ , are similar (and by now routine).

It is interesting to vary this problem by taking the radius  $r_0$  to infinity — in other words, solving Laplace's equation in the whole plane, described in polar coordinates. Then the series of Bessel functions goes over into an integral transform, analogous to the Fourier transform. The initial-data formulas above become

$$g_1(r, \theta) = \sum_{n=0}^{\infty} \int_0^{\infty} d\omega J_n(\omega r) [C_n(\omega) \cos(n\theta) \sinh(\omega L) + D_n(\omega) \sin(n\theta) \sinh(\omega L)],$$

$$C_n(\omega) = \frac{\omega}{\pi} \int_0^{\infty} r dr \int_0^{2\pi} d\theta J_n(\omega r) \cos(n\theta) g_1(r, \theta).$$

(This is not supposed to be obvious; proving it is beyond the scope of this course.)

To clarify the crux of these Bessel expansions, let's strip away the angular complications and summarize them as one-dimensional eigenfunction expansions. Consider an arbitrary function  $f(r)$ .

1. **Fourier–Bessel series:** If the domain of  $f$  is  $0 < r < r_0$ , and  $\omega_{nk} \equiv z_{nk}/r_0$ , then (for a fixed  $n$ )

$$f(r) = \sum_{k=1}^{\infty} A_k J_n(\omega_{nk}r),$$

where

$$A_k = \frac{\int_0^{r_0} J_n(\omega_{nk}r) f(r) r dr}{\int_0^{r_0} J_n(\omega_{nk}r)^2 r dr}.$$

This is a generalization of the Fourier sine series, where the ordinary differential equation involved is a variable-coefficient equation (Bessel's) instead of  $X'' = -\omega^2 X$ .

2. **Hankel transform:** If the domain of  $f$  is  $0 < r < \infty$ , then (for a fixed  $n$ )

$$f(r) = \int_0^\infty A(\omega) J_n(\omega r) \omega d\omega,$$

where

$$A(\omega) = \int_0^\infty f(r) J_n(\omega r) r dr.$$

This is a generalization of the Fourier sine transform.

*Problem 2:* In the  $u_2$  problem the homogeneous boundary conditions are  $Z(0) = 0 = Z(L)$ . This goes with the ODE  $Z'' - \omega^2 Z = 0$ . We see that  $\omega^2$  must be negative this time, so we should change the notation accordingly:

$$\omega^2 \equiv -\nu^2 < 0.$$

We can write the solutions as

$$Z(z) = A \cos(\nu z) + B \sin(\nu z), \quad \nu = \frac{m\pi}{L}.$$

From the  $\theta$  equation (whose boundary conditions are unchanged) we still have  $\mu = n$ . Therefore, the radial equation is

$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left( -\nu^2 - \frac{\mu^2}{r^2} \right) R = 0$$

with  $\mu$  and  $\nu$  related to integers  $n$  and  $m$  as just described.

The solutions of this equation are *modified Bessel functions*, which are regular Bessel functions evaluated at *imaginary argument*. Letting  $\zeta \equiv i\nu r$  puts the equation into standard form:

$$\frac{d^2 R}{d\zeta^2} + \frac{1}{\zeta} \frac{dR}{d\zeta} + \left( 1 - \frac{n^2}{\zeta^2} \right) R = 0.$$

Thus  $R$  as a function of  $\zeta$  is a standard Bessel function, so  $R$  as a function of  $r$  is a modified Bessel function. In the standard notation for modified Bessel functions introduced earlier,  $R$  must be a linear combination of  $I_n(\nu r)$  and  $K_n(\nu r)$ , where  $I$

is the one that is nice at 0 and  $K$  is the one that is nice at infinity. In our problem, zero is the relevant boundary, so  $R(r) = I_n(\nu r)$  and

$$u_2(r, \theta, z) = \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} \sin \frac{m\pi z}{L} [A_{mn} \cos(n\theta) + B_{mn} \sin(n\theta)] I_n \left( \frac{m\pi r}{L} \right).$$

Apply the nonhomogeneous boundary condition:

$$\begin{aligned} f(\theta, z) &= u_2(r_0, \theta, z) \\ &= \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} \sin \frac{m\pi z}{L} [A_{mn} \cos(n\theta) + B_{mn} \sin(n\theta)] I_n \left( \frac{m\pi r_0}{L} \right), \end{aligned}$$

and the coefficients are found by two steps of ordinary Fourier series inversion. In this case the Bessel functions are not used as elements of a basis of eigenfunctions to expand data; rather, they play the same auxiliary role as the  $\sinh(\omega L)$  in some of our Cartesian potential problems and the  $e^{-\omega^2 t} \Big|_{t=0} = 1$  in heat problems.



# Spherical Coordinates and Legendre Functions

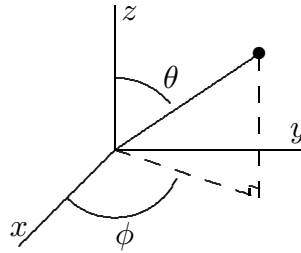
## SPHERICAL COORDINATES

Let's adopt the notation for spherical coordinates that is standard in physics:

$\phi$  = longitude or azimuth,

$\theta$  = colatitude ( $\frac{\pi}{2}$  - latitude) or polar angle.

$$\begin{aligned}x &= r \sin \theta \cos \phi, \\y &= r \sin \theta \sin \phi, \\z &= r \cos \theta.\end{aligned}$$



The ranges of the variables are:  $0 < r < \infty$ ,  $0 < \theta < \pi$ , and  $\phi$  is a periodic coordinate with period  $2\pi$ .

The Laplacian operator is found to be

$$\nabla^2 u = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2}.$$

The term with the  $r$ -derivatives can also be written

$$\frac{1}{r} \frac{\partial^2}{\partial r^2} (ru) \quad \text{or} \quad \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r}.$$

As usual, we try to separate variables by writing

$$u_{\text{sep}} = R(r)\Theta(\theta)\Phi(\phi).$$

We get

$$\frac{r^2 \nabla^2 u}{u} = \frac{(r^2 R')'}{R} + \frac{1}{\sin \theta} \frac{(\sin \theta \Theta')'}{\Theta} + \frac{1}{\sin^2 \theta} \frac{\Phi''}{\Phi}.$$

(Here the primes in the first term indicate derivatives with respect to  $r$ , those in the second term derivatives with respect to  $\theta$ , etc. There is no ambiguity, since each function depends on only one variable.) We have arranged things so that the first

term depends only on  $r$ , and the others depend on  $r$  not at all. Therefore, we can introduce a separation constant (eigenvalue) into Laplace's equation:

$$-\frac{(r^2 R')'}{R} = -K = \frac{1}{\sin \theta} \frac{(\sin \theta \Theta')'}{\Theta} + \frac{1}{\sin^2 \theta} \frac{\Phi''}{\Phi}.$$

Put the  $r$  equation aside for later study. The other equation is

$$\frac{\sin \theta (\sin \theta \Theta')'}{\Theta} + K \sin^2 \theta + \frac{\Phi''}{\Phi} = 0.$$

We can introduce a second separation constant:

$$-\frac{\Phi''}{\Phi} = m^2 = \frac{\sin \theta (\sin \theta \Theta')'}{\Theta} + K \sin^2 \theta.$$

*Remark:* In quantum mechanics,  $K$  has the physical interpretation of the square of the total angular momentum of a particle, while  $m$  is the component of angular momentum about the  $z$  axis.

Just as in two dimensions, problems involving the whole sphere will be different from those involving just a sector. If the region involves a complete sphere, then  $\Phi(\phi)$  must be  $2\pi$ -periodic. Therefore,  $m$  is an integer, and  $\Phi$  is  $A \cos(m\phi) + B \sin(m\phi)$  (or  $C_+ e^{i\phi} + C_- e^{-i\phi}$ ). Then we can write the  $\theta$  equation as

$$\frac{1}{\sin \theta} (\sin \theta \Theta')' + \left[ K - \frac{m^2}{\sin^2 \theta} \right] \Theta = 0.$$

This is an eigenvalue problem for  $K$ . Recall that the proper interval (for the whole sphere) is  $0 < \theta < \pi$ . We have a Sturm–Liouville problem, singular at both endpoints, 0 and  $\pi$ , with weight function  $r(\theta) = \sin \theta$ .

Introduce a new variable by  $x \equiv \cos \theta$  and  $\Theta(\theta) \equiv Z(x) = Z(\cos \theta)$ . (This is not the same as the Cartesian coordinate  $x$ .) Then the equation transforms to the purely algebraic form

$$(1 - x^2) \frac{d^2 Z}{dx^2} - 2x \frac{dZ}{dx} + \left[ K - \frac{m^2}{1 - x^2} \right] Z = 0$$

on the interval  $-1 < x < 1$ . The first two terms can be combined into

$$\frac{d}{dx} \left[ (1 - x^2) \frac{dZ}{dx} \right].$$

Since  $dx = -\sin \theta d\theta$ , the weight factor is now unity.

If  $m = 0$ , this equation is called *Legendre's equation* and the solutions are *Legendre functions*. Solutions of the equation with  $m \neq 0$  are *associated Legendre functions*.

We concentrate first on  $m = 0$ . (This means that we are looking only at solutions of the original PDE that are rotationally symmetric about the  $z$  axis — i.e., independent of  $\phi$ .) We now get the payoff from a problem that you may have studied in differential equations or in linear algebra or both. When the equation is solved by power series (method of Frobenius), one finds that if  $K = l(l + 1)$ , where  $l$  is a nonnegative integer, then there is one solution (of the two independent ones) that is a *polynomial* — the Frobenius series terminates. These are called the *Legendre polynomials*,  $P_l(x)$ , and a totally different way of stumbling upon them is to apply the Gram–Schmidt orthogonalization procedure to the sequence of powers,  $\{1, x, x^2, \dots\}$ , regarded as functions on the interval  $[-1, 1]$  with the usual inner product. The first few of them (normalized so that  $P(\cos 0) = P(1) = 1$ ) are

$$P_0(x) = 1$$

$$P_1(x) = x; \quad \Theta_1(\theta) = \cos \theta$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1); \quad \Theta_2(\theta) = \frac{1}{2}(3 \cos^2 \theta - 1)$$

$P_l(x)$  is a polynomial of degree  $l$ . It is given explicitly by *Rodrigues's formula*,

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l.$$

Just as we required solutions in polar coordinates to be bounded at the origin, we must require solutions in spherical coordinates to be bounded at the north and south poles ( $x = \pm 1$ ). It is a fact that all solutions except the polynomials  $P_l$  behave unacceptably at one or the other of the endpoints. In our problem, therefore, the eigenvalues are the numbers  $l(l + 1)$ , and the Legendre polynomials are the eigenvectors. The other solutions become relevant in other PDE problems where the region does not contain the whole sphere (a cone, for instance). When  $K = l(l + 1)$  (so that  $P_l$  exists), another, linearly independent, solution can be found by the method of *reduction of order* or the general Frobenius theory [review your ODE textbook]. It is called  $Q_l$ .

$$Q_0(x) = \frac{1}{2} \ln \left( \frac{1+x}{1-x} \right), \quad Q_1(x) = \frac{x}{2} \ln \left( \frac{1+x}{1-x} \right) - 1.$$

It's clear that any linear combination of  $P$  and  $Q$  with a nonzero  $Q$  component is singular at the endpoints.

The orthogonality and normalization properties of the Legendre polynomials are

$$\int_{-1}^1 P_l(x) P_k(x) dx = 0 \quad \text{if } l \neq k,$$

$$\int_{-1}^1 P_l(x)^2 dx = \frac{2}{2l+1}.$$

Note that  $\int_{-1}^1 [\dots x \dots] dx$  is the same as  $\int_0^\pi [\dots \cos \theta \dots] \sin \theta d\theta$ . The factor  $\sin \theta$  is to be expected on geometrical grounds; it appears naturally in the volume element in spherical coordinates,

$$dV = dx dy dz = r^2 \sin \theta dr d\theta d\phi,$$

and the surface area element on a sphere,

$$dS = r_0^2 \sin \theta d\theta d\phi.$$

Now let's return to the radial equation,

$$r(rR)'' = l(l+1)R,$$

that came out of Laplace's equation. Its solutions are

$$R(r) = Ar^l + Br^{-l-1}.$$

(Except for the  $-1$  in the second exponent, this is just like the two-dimensional case.) We note that one of the basis solutions vanishes as  $r \rightarrow 0$ , the other as  $r \rightarrow \infty$ .

Now we can put all the pieces together to solve a boundary value problem with no  $\phi$  dependence. (If the problem has this axial symmetry and the solution is unique, then the solution must also have that symmetry. Clearly, this will require axially symmetric boundary data.) If the region in question is a ball (the interior of a sphere), then the form of the general axially symmetric solution is

$$u(r, \theta) = \sum_{l=0}^{\infty} b_l r^l P_l(\cos \theta).$$

If Dirichlet boundary data are given on the sphere, then

$$f(\theta) \equiv u(r_0, \theta) = \sum_{l=0}^{\infty} b_l r_0^l P_l(\cos \theta)$$

for all  $\theta$  between 0 and  $\pi$ . Therefore, by the orthogonality and normalization formulas previously stated,

$$b_l = \frac{2l+1}{2r_0^l} \int_0^\pi f(\theta) P_l(\cos \theta) \sin \theta d\theta.$$

If the region is the exterior of a sphere, we would use  $r^{-(l+1)}$  instead of  $r^l$ . For the shell between two spheres, we would use both, and would need data on both surfaces to determine the coefficients. As always, Neumann or Robin data instead of Dirichlet might be appropriate, depending on the physics of the individual problem.

## SPHERICAL HARMONICS

What if the boundary data do depend on  $\phi$  as well as  $\theta$ ? We must generalize the sum to

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l b_{lm} r^l P_l^m(\cos \theta) e^{im\phi},$$

where the functions  $P_l^m$ , called *associated Legendre functions*, are solutions of

$$[(1-x^2)P']' + \left[ l(l+1) - \frac{m^2}{1-x^2} \right] P = 0.$$

The condition of regularity at the poles forces  $|m| \leq l$ , and this constraint has been taken into account by writing the sum over  $m$  from  $-l$  to  $l$ . There is a generalized Rodrigues formula,

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l.$$

These provide a *complete, orthogonal set of functions* on (the surface of) a sphere. The basis functions most commonly used are called *spherical harmonics*, defined by

$$Y_l^m(\theta, \phi) = \left[ \frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{\frac{1}{2}} P_l^m(\cos \theta) e^{im\phi}$$

for  $-l < m < l$  and  $l = 0, 1, \dots$ . The purpose of the complicated numerical coefficient is to make them orthonormal. Integration over the sphere is done with respect to the usual area element,

$$\int \dots d\Omega \equiv \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \dots$$

Then one has the orthonormality relation

$$\int d\Omega Y_{l'}^{m'}(\theta, \phi)^* Y_l^m(\theta, \phi) = \begin{cases} 1 & \text{if } l' = l \text{ and } m' = m, \\ 0 & \text{otherwise.} \end{cases}$$

The completeness (basis) property is: An arbitrary\* function on the sphere (i.e., a function of  $\theta$  and  $\phi$  as they range through their standard intervals) can be expanded as

$$g(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm} Y_l^m(\theta, \phi),$$

---

\* sufficiently well-behaved, say square-integrable

## A table of the first few spherical harmonics

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}$$

$$Y_1^1 = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi}$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos \theta$$

$$Y_1^{-1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi}$$

$$Y_2^2 = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\phi}$$

$$Y_2^1 = -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi}$$

$$Y_2^0 = \sqrt{\frac{5}{4\pi}} \left( \frac{3}{2} \cos^2 \theta - \frac{1}{2} \right)$$

$$Y_2^{-1} = \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{-i\phi}$$

$$Y_2^{-2} = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{-2i\phi}$$

where

$$A_{lm} = \int d\Omega Y_l^m(\theta, \phi)^* g(\theta, \phi).$$

This, of course, is precisely what we need to solve the potential equation with arbitrary boundary data on a spherical boundary. But such a way of decomposing functions on a sphere may be useful even when no PDE is involved, just as the Fourier series and Fourier transform have many applications outside differential equations. For example, the shape of the earth (as measured by the gravitational attraction on satellites) is represented by a sum of spherical harmonics, where the first (constant) term is by far the largest (since the earth is nearly round). The three terms with  $l = 1$  can be removed by moving the origin of coordinates to the right spot; this defines the “center” of a nonspherical earth. Thus the first interesting terms are the five with  $l = 2$ ; their nonzero presence is called the *quadrupole moment* of the earth. Similar remarks apply to the analysis of any approximately spherical object, force field, etc.\*

\* See, for example, M. T. Zuber et al., “The Shape of 433 Eros from the NEAR-Shoemaker Laser Rangefinder,” *Science* **289**, 2097–2101 (2000), and adjacent articles, for

A sensible person does not try to memorize all the formulas about spherical harmonics (or any other class of special functions). The point is to understand that they exist and why they are useful. The details when needed are looked up in handbooks or obtained from computer software. Complicated formulas should not obscure the beauty and power of our march from a basis of eigenvectors in  $\mathbf{R}^2$ , through Fourier series in one dimension, to this basis of eigenfunctions on a sphere!

## SPHERICAL BESSEL FUNCTIONS

Instead of the potential equation,  $\nabla^2 u = 0$ , consider now the Helmholtz equation,

$$\nabla^2 u = -\omega^2 u.$$

This will arise from the separation of variables in the wave or heat equation in three dimensions. When we continue the separation in spherical coordinates, the angular part is exactly the same as before, so the angular dependence of solutions of the Helmholtz equation is still given by the spherical harmonics (or Legendre polynomials, in the axially symmetric case). The radial equation, however, becomes

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left[ \omega^2 - \frac{l(l+1)}{r^2} \right] R = 0.$$

Thus the radial solutions are no longer just powers of  $r$ .

Let  $z \equiv \omega r$ ,  $Z(z) \equiv \sqrt{z} R$ . Then (another exercise)

$$\frac{d^2 Z}{dz^2} + \frac{1}{z} \frac{dZ}{dz} + \left[ 1 - \frac{l(l+1) + \frac{1}{4}}{z^2} \right] Z = 0.$$

This is Bessel's equation, with  $\mu = l + \frac{1}{2}$  (since  $(l + \frac{1}{2})^2 = l(l+1) + \frac{1}{4}$ ). The consequent solutions

$$R(r) = \frac{1}{\sqrt{\omega r}} J_{l+\frac{1}{2}}(\omega r)$$

are called *spherical Bessel functions*, with the notation

$$j_l(z) \equiv \sqrt{\frac{\pi}{2z}} J_{l+\frac{1}{2}}(z).$$

Similarly, the other types of Bessel functions have their spherical counterparts,  $y_l$ ,  $h_l^{(1)}$ , etc.

---

an analysis of a potato-shaped asteroid. There the harmonics with factors  $e^{\pm im\phi}$  are combined into real functions with factors  $\cos m\phi$  and  $\sin m\phi$ , so the five coefficients for  $l = 2$  are named  $C_{20}$ ,  $C_{21}$ ,  $S_{21}$ ,  $C_{22}$ ,  $S_{22}$ .

The surprising good news is that these fractional-order Bessel functions are not an entirely new family of functions. They can all be expressed in terms of sines and cosines. One has

$$j_0(z) = \frac{\sin z}{z}, \quad y_0(z) = -\frac{\cos z}{z}$$

(note that  $j_0$  is regular at 0 and  $y_0$  is not, as expected from their definitions),

$$j_1(z) = \frac{\sin z}{z^2} - \frac{\cos z}{z},$$

and, in general,

$$j_l(z) = z^l \left( -\frac{1}{z} \frac{d}{dz} \right)^l \frac{\sin z}{z},$$

$$y_l(z) = -z^l \left( -\frac{1}{z} \frac{d}{dz} \right)^l \frac{\cos z}{z}.$$

Notice that for large  $l$  they contain many terms, if all the derivatives are worked out.

We would naturally want to use these to solve a PDE with a homogeneous boundary condition on a sphere. As in the case of integer-order Bessel functions, there will be a normal mode corresponding to each value of  $z$  for which  $j_l(z)$  vanishes (or its derivative vanishes, if the boundary condition is of the Neumann type). To find these roots one needs to solve a trigonometric equation, as in the classic Sturm–Liouville problems; many of the small roots can be looked up in tables, and there are approximate asymptotic formulas for the large ones. The resulting normal modes form a complete, orthogonal set for expanding functions in the interior of a ball.



## Classification of Second-Order Linear Equations

We have looked at three fundamental partial differential equations:

$$\text{Laplace: } \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$$\text{wave: } \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$$

$$\text{heat: } \frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} = 0.$$

Each of these turned out to have its own characteristic properties, which we want to review here and put in a more general context. Of particular interest for each equation are

- (1) what sort of data (initial or boundary conditions) are needed to constitute a *well-posed problem* — one with exactly one solution;
- (2) smoothness of the solutions;
- (3) how the influence of the data spreads (*causality* or *finite propagation speed*).

The most general *second-order linear* differential equation in two variables, say  $x$  and  $y$ , looks like

$$L[u] \equiv A(x, y) \frac{\partial^2 u}{\partial x^2} + B(x, y) \frac{\partial^2 u}{\partial x \partial y} + C(x, y) \frac{\partial^2 u}{\partial y^2} \\ + D(x, y) \frac{\partial u}{\partial x} + E(x, y) \frac{\partial u}{\partial y} + F(x, y)u = 0,$$

where  $A, \dots, F$  are functions of  $x$  and  $y$ . Suppose just for a moment that these coefficients are *constants*. Then the long expression is reminiscent of the formula for the most general *conic section*. Indeed, if we replace each  $\partial/\partial x$  by a new variable,  $X$ , and replace each  $\partial/\partial y$  by  $Y$ , and replace  $L$  by  $0$ , then we get exactly the conic section equation:

$$0 = AX^2 + BXY + CY^2 + DX + EY + F.$$

Now recall from analytic geometry that it is always possible to make a *rotation* of axes in the  $X$ - $Y$  space after which the cross-term coefficient  $B$  is zero. Suppose that this has been done:

$$0 = AX^2 + CY^2 + DX + EY + F.$$

Then recall that (if certain “degenerate cases” are ignored) the curve described by this equation is an

ellipse if  $A$  and  $C$  have the same sign,  
 hyperbola if  $A$  and  $C$  have opposite signs,  
 parabola if one of them ( $A$  or  $C$ ) is 0.

We assign the same terminology to the partial differential equations that result when  $X$  is replaced by  $\partial/\partial x$ , etc. Thus Laplace’s equation is elliptic, the wave equation is hyperbolic, and the heat equation is parabolic. (In the latter two cases  $y$  is called  $t$  for physical reasons.)

Now suppose that  $A$ , etc., do depend on  $x$  and  $y$ . Then at each point  $(x, y)$  it is possible to find a rotation

$$\begin{aligned}\frac{\partial}{\partial x'} &= \cos \theta \frac{\partial}{\partial x} - \sin \theta \frac{\partial}{\partial y}, \\ \frac{\partial}{\partial y'} &= \sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y},\end{aligned}$$

which eliminates the  $B(x, y)$  term. (The angle  $\theta$  may depend on  $x$  and  $y$ , so  $B$  is not necessarily zero at other points.) The character of the PDE at that point is defined to be elliptic, hyperbolic, or parabolic depending on the signs of the coefficients of the new coefficients  $A$  and  $C$  there. The *discriminant*

$$\Delta \equiv B(x, y)^2 - 4A(x, y)C(x, y)$$

is not changed by a rotation of coordinates. Therefore, it is easy to see that the equation is

elliptic if  $\Delta < 0$ ,  
 hyperbolic if  $\Delta > 0$ ,  
 parabolic if  $\Delta = 0$ .

For most equations of practical interest, the operator will be of the same type at all points.

The classification can be extended to *nonlinear* equations, provided they are *linear in their dependence on the second derivatives of  $u$* . Such an equation is called *quasilinear*. Example:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + u^3 = 0$$

is quasilinear and elliptic.

*Remark:* From linear algebra you may recall that what we are doing here is diagonalizing the matrix (quadratic form)

$$\begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix},$$

that the *new*  $A$  and  $C$  are the eigenvalues of that matrix, and that  $\Delta$  is  $-4$  times its determinant. This is the secret to extending the classification to equations in more than two variables, such as

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0.$$

This example counts as hyperbolic, since it has one coefficient with sign opposite to the others. More generally, there is a coefficient matrix which has to be diagonalized, and the signs of its eigenvalues are what counts: The operator is elliptic if all the signs are the same, hyperbolic if one is different, and parabolic if one eigenvalue is zero and the others have the same sign. (There are other possibilities, such as two positive and two negative eigenvalues, but they seldom arise in applications.)

Now let's discuss the three matters listed at the beginning. The facts I'm about to state are generalizations of things we already know about the heat, wave, and Laplace equation.

(1) In a hyperbolic or parabolic equation, we identify the "special" coordinate as the *time*. (That is the coordinate with the strange sign in the hyperbolic case or the zero in the parabolic case. In the latter case we assume that the first-order derivative with respect to  $t$  does occur (multiplied by a real coefficient), although by hypothesis the second-order one doesn't.) Then the fact is that these equations behave just like ordinary differential equations as to initial data: The parabolic equation is first-order, so a solution is uniquely determined by its initial value,  $u(x, \dots, t = 0)$ . The hyperbolic equation is second-order, so you need also the initial value of the time derivative. Boundary conditions at the edges of the spatial domain (let's call it  $D$ ) may also be necessary to specify the solution, as we well know from examples. (These latter boundary conditions are of the same type as needed to produce a well-posed problem for an *elliptic* equation on  $D$  — see below. This is not surprising, since the spatial equation we get when a hyperbolic or parabolic equation (in 3 or more variables) is solved by separation of variables is an elliptic equation, such as  $\nabla^2 \phi = -\omega^2 \phi$ .)

In the parabolic case, a solution is guaranteed to exist only in one direction of time from the initial data surface. (With the usual choice of signs, this is the positive time direction.) If you try to solve the heat equation in the negative direction, a solution may not exist for the given data; when solutions do exist, they are *unstable* in the sense that a small change in the data creates drastic changes in the solution.

Since real science and engineering deal with only approximately measured data, this makes the solution in the backward direction almost useless in practice.

For an elliptic equation, one might expect to have a well-posed problem given the value of  $u$  and its normal derivative on an “initial” surface, since the equation is second-order in every variable. However, it turns out that a solution may not exist for all data; it will exist in a neighborhood of the surface, but it will “blow up” somewhere else. When solutions exist, they may be unstable. Instead, the proper and natural boundary condition for an elliptic equation (as we know from physical applications of Laplace’s equation) is to prescribe the function or its derivative (but not both) at every point on a *closed curve or surface* surrounding a region. (Conversely, this sort of boundary condition will *not* give a well-posed problem for a hyperbolic or parabolic equation.)

I have been using the term *well-posed* without formally defining it. It means, above all, that the problem (consisting, typically, of a differential equation plus boundary conditions) has been stated so that it has exactly one solution. Stating too few conditions will make the solution nonunique; too many conditions, and it will not exist; try to use the wrong kind of conditions (e.g., initial data for an elliptic equation), and there will be no happy medium! In addition, it is customary to require *stability*; that is, that the solution depends continuously on the data.

(2) Elliptic and parabolic equations (with smooth coefficients) have solutions that are *smooth* (that is, differentiable arbitrarily many times), regardless of how rough their data (boundary values) are. But solutions of hyperbolic equations may be nondifferentiable, discontinuous, or even distributions — such as  $\delta(x - ct)$  for the wave equation. In other words, singularities in the initial data are propagated by a hyperbolic equation into the solution region.

(3) Hyperbolic equations spread the initial data out into space at a finite “wave” speed. (In applications, this is the speed of sound, the speed of light, etc.) In contrast, the initial data of the heat equation can instantly affect the solution arbitrarily far away.

There is one more type of second-order linear equation: the (*time-dependent*) *Schrödinger equations* of quantum mechanics, of which the simplest case is

$$i \frac{\partial u}{\partial t} = - \frac{\partial^2 u}{\partial x^2}.$$

This was overlooked in the classification above, because we were tacitly assuming that all the quantities were real. The Schrödinger equation does not fit into any of the three standard categories; instead, it shares some of the features of parabolic and hyperbolic equations.

- Like the heat equation, it is first-order in time. Therefore,  $u(x, 0)$  (by itself) is appropriate initial data.

- Unlike the heat equation, but like the wave equation, its solutions are not necessarily smooth. Unlike the wave equation, the singularities in the solutions can disappear and then reappear at later times; this happens most notoriously for the Green function of the *harmonic oscillator* equation

$$i \frac{\partial u}{\partial t} = - \frac{\partial^2 u}{\partial x^2} + x^2 u,$$

which contains the periodic factor  $\csc(2t)$ .

- Unlike the wave equation, but like the heat equation, its solutions are not limited by a finite propagation speed.
- Its nicest property is *unitarity*: The  $L^2$  norm of the solution at any fixed  $t$  is the same as the  $L^2$  norm of the initial data. That is,

$$\int_D |u(x, t)|^2 dx = \int_D |u(x, 0)|^2 dx.$$

(Here it is assumed that the differential operator in the spatial variables (the Hamiltonian) is self-adjoint.)

#### THE MAXIMUM PRINCIPLE

Consider an *elliptic* or *parabolic* PDE (second-order, linear, homogeneous),

$$A(x, y) \frac{\partial^2 u}{\partial x^2} + B(x, y) \frac{\partial^2 u}{\partial x \partial y} + C(x, y) \frac{\partial^2 u}{\partial y^2} + D(x, y) \frac{\partial u}{\partial z} + E(x, y) \frac{\partial u}{\partial y} + F(x, y)u = \epsilon \frac{\partial u}{\partial t}$$

where  $\epsilon = 0$  or  $1$ . (In this discussion  $u$  is a function of two variables or of three, depending on the context, but we'll usually suppress  $t$  in the notation.) If a solution has a maximum or minimum, it occurs either on the *boundary* of the region considered or at a *critical point* (where all first-order partial derivatives of  $u$  are zero). Consider the latter possibility. We may assume for simplicity that the critical point is at the origin and that  $B(0, 0)$  has been eliminated by a rigid rotation of the  $x$  and  $y$  axes. So, at the origin the equation reduces to

$$A(0, 0) \frac{\partial^2 u}{\partial x^2} + C(0, 0) \frac{\partial^2 u}{\partial y^2} + F(0, 0)u = 0,$$

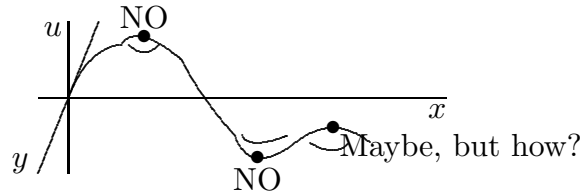
where  $A$  and  $C$  are *both positive*, by definition of elliptic or parabolic.

If  $F(0, 0)$  is negative and  $u(0, 0)$  is positive, then

$$A(0, 0) \frac{\partial^2 u}{\partial x^2}(0, 0) + C(0, 0) \frac{\partial^2 u}{\partial y^2}(0, 0) > 0,$$

so at least one of the second-order partials is positive. Therefore,  $u(0,0)$  *cannot be a local maximum*. Similarly, if  $F(0,0)$  and  $u(0,0)$  are both negative, then at least one of the second-order partials is negative, so  $u(0,0)$  *cannot be a local minimum*. Putting these facts together, we can conclude:

**Theorem (a simple maximum principle):** If  $F(x,y)$  is always negative, then  $u$  cannot have a positive maximum nor a negative minimum in the interior of its domain.



We have seen this principle at work in the eigenvalue problem for  $\nabla^2$ . If  $-\nabla^2 u = \lambda u$ , then we found that  $\lambda$  must be positive for Dirichlet boundary conditions and nonnegative for Neumann boundary conditions; thus  $F$  is nonnegative and the theorem does *not* apply. That allows us to have eigenfunctions like those pictured on pp. 43 and 113, which blatantly violate the conclusion of the theorem. On the other hand, when  $\lambda < 0$ , or  $\lambda = 0$  and  $u$  is not constant, then the theorem applies; indeed, the solutions (of the PDE) that we found in such cases were always *concave away* from the coordinate plane in at least one dimension (like  $e^x$ ), and that was why we could never find solutions for such  $\lambda$  that satisfied all the homogeneous boundary conditions needed to be eigenfunctions.

A somewhat more technical proof yields a somewhat simpler sounding theorem:

**Theorem (a maximum principle):** If  $F(x,y) = 0$  everywhere, then  $u$  cannot have an interior local extremum of *either* kind, except in the trivial case where  $u$  is a constant function.

For the special case of Laplace's equation,  $\nabla^2 u = 0$ , this maximum principle follows from the theorem that

$$u(\mathbf{x}) = (\text{average of } u \text{ over a circle centered at } \mathbf{x}),$$

which in turn is easy to see from the expansion of  $u$  in a Fourier series (in the polar angle) inside the circle (see p. 103). The same thing is true for Laplace's equation in 3 dimensions, with the circle replaced by a sphere and the Fourier series by the expansion in spherical harmonics.

As we've seen, the maximum principle holds only for a rather restricted class of differential equations: not only must the equation be elliptic or parabolic, but also

there is a sign condition on the terms without derivatives. Nevertheless, various forms of the maximum principle are important tools in proving theorems about the properties of solutions. Here are two examples:

**Corollary 1:** if  $F = 0$  everywhere and the domain of  $u$  is bounded, then the global maximum and minimum values of  $u$  occur on the boundary.

**Corollary 2:** If  $F \leq 0$  everywhere, then  $u = 0$  everywhere on the boundary implies that  $u = 0$  everywhere.

**COROLLARY TO THE COROLLARY: Uniqueness Theorem:** If a homogeneous linear problem satisfies the maximum principle in the sense of Corollary 2, then the solution of an associated nonhomogeneous problem is unique.

**EXAMPLE:** Consider the Poisson equation  $-\nabla^2 u = j$  in a bounded region with the nonhomogeneous boundary condition that  $u(\mathbf{x}) = f(\mathbf{x})$  for all  $\mathbf{x}$  on the boundary of the region. (The functions  $j$  and  $f$  are fixed, as part of the statement of the problem.) Then the solution (if we assume that it exists) is unique: If there were two of them,  $u_1$  and  $u_2$ , then  $v \equiv u_1 - u_2$  would satisfy  $\nabla^2 v = 0$  in the region and  $v = 0$  on the boundary, so  $v$  would be identically zero, a contradiction.

## Appendix A

### The Heat Equation

The *heat equation* or *diffusion equation* in one space dimension is

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}. \quad (*)$$

It's a *partial differential equation* (PDE) because partial derivatives of the unknown function with respect to two (or more) variables appear in it.

Here is a brief survey of the physics of the problem:

**1. Physical interpretation and derivation:** In the most usual application,  $t$  is time, and  $x$  is a spatial variable (say along a thin wire, a homogeneous bar, or an imaginary one-dimensional world studied in the (justified) hope that the solutions of the more difficult three-dimensional heat equation

$$\nabla^2 u = \frac{\partial u}{\partial t}$$

will be qualitatively similar).  $u(t, x)$  is the temperature in the bar (possibly with something subtracted off, as we'll see). The equation follows quickly from algebraic formulations of the physical principles that

- (1) the amount of heat energy in any small region of the bar is proportional to the temperature there,
- (2) the rate of heat flow is proportional to the derivative of the temperature, since it's driven by temperature differences between regions.

In fact, the same equation describes many other *diffusion* processes. It — or some modification of it — arises whenever one studies the large-scale, averaged effects of the random motion of many particles. (Think of a cloud of mosquitos released from a cage in one corner of a large room.)

**2. Scaling to remove irrelevant constants:** We are free to redefine the units in which  $u$ ,  $t$ , and  $x$  are measured. In general, the equation will first be presented to us as

$$K \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t},$$

where  $K$  is a constant depending on the physical properties of the material (specifically, its specific heat and thermal conductivity, which are the coefficients in the two “algebraic formulations” mentioned above). By rescaling  $x$  or  $t$  (or both), we



can change  $K$  to 1. So there is no loss of generality in ignoring  $K$  henceforth. This uses up only one of the three degrees of freedom in the units. The other two can be used in other ways.

Typically, our bar will have a finite length, say  $L$ . We can rescale  $x$  to make  $L$  have any convenient value; the most popular choices are 1 (not surprisingly) and  $\pi$  (for reasons that will become obvious later). After that, we can rescale  $t$  so as to keep  $K$  equal to 1. We can also add a constant to  $x$  so that the left endpoint of the bar is at  $x = 0$ .

Scaling  $u$  will not change the form of the equation, since it is linear (see below). However, this scaling freedom can be used to simplify a boundary condition or initial condition.

**3. Initial and boundary conditions:** To make a PDE into a well-defined problem, we have to state over what domain of the independent variables we hope to solve it, and we need to have enough information about the behavior of  $u$  on the boundary of that domain to make the solution of the problem unique. For physical and mathematical reasons, time and space enter the heat problem in different ways. One finds:

- (1) If we know the temperature distribution at one time (say  $t = 0$ ), we can hope to predict the temperature at *later* times, but not necessarily at *earlier* times. (If we observe a room full of mosquitos, it is hard to tell by looking which corner they flew out of.) Thus we will be solving (\*) in the region

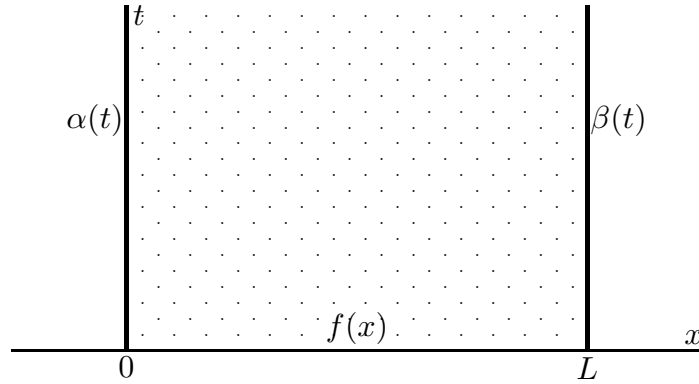
$$0 < x < L, \quad t > 0$$

given *initial data*

$$u(0, x) = f(x) \quad \text{for } 0 < x < L.$$

- (2) We need to know what happens to the heat when it reaches the end of the bar. Obviously it will make a big difference to the temperature distribution whether the end is insulated or in contact with some other material which can conduct heat away. There are four standard types of *boundary conditions* that can be considered. Each type is worthy of consideration for its own sake as a mathematical possibility, but it happens that each one has a real physical interpretation in the heat problem:

- (A) *Dirichlet* condition:  $u(t, 0) = \alpha(t)$  for some given function  $\alpha$ . This says that the temperature at the end of the bar is controlled (say by contact with a “heat bath”).



(B) *Neumann* condition:  $\frac{\partial u}{\partial x}(t, 0) = \alpha(t)$ . This says that the heat flow through the end is controlled. This is hard to do in practice, except in the special case  $\alpha = 0$ , which says that the end is *insulated*.

(C) A generalization of the first two is a *Robin* condition:

$$c_1 u(t, 0) + c_2 \frac{\partial u}{\partial x}(t, 0) = \alpha(t),$$

where the  $c$ 's are constants characteristic of the situation. Such a condition arises in *convective cooling*, when the bar is in contact with a less dense medium (such as air) which can carry away heat, but not fast enough to lower the bar temperature immediately to the medium's temperature.

In all these cases of conditions at  $x = 0$ , one would need another condition (not necessarily the same kind) at  $x = L$  to complete the specification of the problem.

(D) *Periodic* boundary conditions: These deal with both endpoints at once.

$$u(t, 0) = u(t, L), \quad \frac{\partial u}{\partial x}(t, 0) = \frac{\partial u}{\partial x}(t, L).$$

The usual physical interpretation of this is that our "bar" is actually a *ring*, and  $x$  is an angle. (Thus  $L = 2\pi$  when  $x$  is measured in radians.)

One tends to think of the boundary conditions as part of the definition of the *physical system* under study, while the initial conditions label the various *possible solutions* of the equations of motion of that given system. In other words, in our discussions the boundary conditions are usually "more constant", the initial conditions "more variable". Imposing the initial conditions is usually the last step in finding a solution, as it is usually is for ODEs, too.

#### THE STEADY-STATE SOLUTION.

We shall now complete the solution of the one-dimensional heat problem with fixed end temperatures (in mathematical terms, nonhomogeneous Dirichlet data

that are independent of time). The overall solution strategy is outlined at the end of the section “Fundamental concepts” in the main text of the notes; we continue from there.

Return to step (1) and assume that  $v(t, x) = V(x)$ . Then the equation becomes  $0 = V''$ , and the boundary conditions become  $V(0) = T_1$ ,  $V(1) = T_2$ . We see that  $V = C_1x + C_2$  and thus

$$T_1 = C_2, \quad T_2 = C_1 + C_2.$$

Therefore,

$$V(x) = (T_2 - T_1)x + T_1.$$

REMARK: Try to repeat this construction for the boundary conditions

$$V'(0) = F_1, \quad V'(1) = F_2$$

(nonhomogeneous *Neumann* data). Something strange happens. Can you figure out what to do now? The two-dimensional generalization of this phenomenon is treated in the main text in the section on “rectangles”.

#### SEPARATION OF VARIABLES

Now return to the second half of the problem, the initial-value problem for the heat equation with homogenized boundary conditions:

PDE: 
$$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2},$$

BC: 
$$w(t, 0) = 0, \quad w(t, 1) = 0,$$

IC: 
$$w(0, x) = g(x) \quad [= f(x) - V(x)].$$

Our strategy will be to look first for functions of the form

$$w_{\text{sep}}(t, x) = T(t)X(x)$$

which satisfy all the *homogeneous* equations of the problem (namely, the PDE and BC) — but *not* (usually) the nonhomogeneous equations (the IC, in this case). Then we will try to satisfy the nonhomogeneous conditions by a “superposition”

(or infinite linear combination) of these separated solutions: It will look something like

$$w(t, x) = \sum_{n=1}^{\infty} c_n T_n(t) X_n(x).$$

At risk of tedium, let me emphasize again that

- (1) since the separated (product) solutions satisfy the homogeneous conditions, the sum will also;
- (2) attempting to impose the nonhomogeneous conditions on the individual  $w_{\text{sep}}$ 's will lead to catastrophe, since nonhomogeneous conditions are not preserved under summation. If we found an infinite string of functions that *each* satisfied the nonhomogeneous condition  $u(t, 0) = T_1$ , then the sum of their boundary values would be an infinite series of equal constants, which would not converge — certainly not to  $T_1$ .

Substitute  $w = TX$  into the PDE:

$$T'(t)X(x) = T(t)X''(x).$$

Now we *separate* the variables: Divide by  $T(t)X(x)$ , getting

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}.$$

In this equation the left side depends only on  $t$  and the right side depends only on  $x$ . The only way the equation can then hold for all  $t$  and all  $x$  is that both quantities are *constant*:

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = -\lambda.$$

(I have advance information that the most interesting values of this constant will be negative, so I call it  $-\lambda$ . However, we are not yet ready to make any commitment as to whether  $\lambda$  is positive, negative, zero, or even complex. All possibilities must be considered.)

We have split the equation into two equations,

$$X'' + \lambda X = 0, \tag{1}$$

$$T' + \lambda T = 0. \tag{2}$$

Now look at the boundary conditions, which are

$$0 = w_{\text{sep}}(t, 0) = T(t)X(0), \quad 0 = w_{\text{sep}}(t, 1) = T(t)X(1).$$

These impose restrictions on  $X$ , not  $T$ . (If we were to satisfy either of them for all  $t$  by setting  $T(t) = 0$ , we would make the entire solution  $w_{\text{sep}}$  identically equal to 0, a trivial and uninteresting solution.) Our next task is to find the values of  $\lambda$  that allow  $X$  to vanish at both 0 and 1.

Suppose first that  $\lambda$  is positive, and write  $\lambda = \omega^2$  (where  $\omega$  is positive). Then (1) with its BC is

$$X'' + \omega^2 X = 0, \quad X(0) = 0, \quad X(1) = 0.$$

The general solution of the ODE is

$$X = c_1 \cos \omega x + c_2 \sin \omega x, \quad (\dagger)$$

and the first boundary condition forces  $c_1 = 0$ . We can choose  $c_2 = 1$  without loss of generality (since what we are looking for is a linearly independent set of separated solutions  $w_{\text{sep}}$ ). So  $X = \sin \omega x$ . Then the second boundary condition is

$$\sin \omega = 0.$$

The positive solutions of this equation are

$$\omega_n \equiv n\pi, \quad n = 1, 2, \dots .$$

[Notice that the root  $\omega = 0$  is irrelevant, since solutions of the ODE with  $\lambda = 0$  do not have the form  $(\dagger)$ . Negative  $\omega$ 's give nothing new, which is why we restricted  $\omega$  to be positive when we introduced it.] Note, incidentally, that if we were working on the interval  $0 < x < \pi$  instead of  $0 < x < 1$ , we would get just  $\omega_n = n$ , without the  $\pi$ .

We can now solve the time equation, (2):

$$T(t) = e^{-\omega_n^2 t} = e^{-n^2 \pi^2 t}.$$

The full separated solution for each  $n$  is thus

$$w_{\text{sep}}(t, x) = \sin(n\pi x) e^{-n^2 \pi^2 t}.$$

Now consider the possibility that  $\lambda = 0$ . In place of  $(\dagger)$  we have the general solution

$$X = c_1 + c_2 x.$$

Applying the two BC, we swiftly get  $c_1 = 0$ ,  $c_2 = 0$ . So there is no nontrivial solution with  $\lambda = 0$ .

Similar arguments show that negative and complex  $\lambda$ 's give only trivial solutions. In the negative case, write  $\lambda = -\kappa^2$ ; then

$$X = c_1 \cosh \kappa x + c_2 \sinh \kappa x,$$

and the result follows (since  $\cosh 0 \neq 0$  and  $\sinh z \neq 0$  unless  $z = 0$ ). If  $\lambda$  is complex, it has two complex square roots, which are negatives (*not* complex conjugates!) of each other. Thus

$$X = c_1 e^{(\kappa+i\omega)x} + c_2 e^{-(\kappa+i\omega)x},$$

where  $(\kappa + i\omega)^2 = -\lambda$  and  $\kappa \neq 0$  (else we would be back in the case of positive  $\lambda$ ).  $X(0) = 0$  implies that  $c_2 = -c_1$ , and then  $X(1) = 0$  implies that

$$e^{(\kappa+i\omega)} = e^{-(\kappa+i\omega)}.$$

Since  $\kappa \neq 0$ , these two complex numbers have different moduli (absolute values), so this conclusion is a contradiction.

There is a more modern, less grubby way to see that  $\lambda$  has to be positive. Using the ODE ( $X'' = -\lambda X$ ) and the BC (which allow us to discard all endpoint terms which arise in integration by parts), we see that

$$\begin{aligned} \lambda \int_0^1 |X(x)|^2 dx &= - \int_0^1 X^* X'' dx \\ &= + \int_0^1 |X'|^2 dx \\ &= - \int_0^1 (X'')^* X dx \\ &= +\lambda^* \int_0^1 |X|^2 dx. \end{aligned}$$

Comparing the first and last members of this chain of equalities, we see that  $\lambda = \lambda^*$  — that is,  $\lambda$  must be real. Comparing either of the extreme members with the one in the middle, we see that  $\lambda$  is positive, since two integrals are positive.

This argument suggests a general method for handling such questions when the second-derivative operator is replaced by a more general linear differential operator  $L[X]$ . If the  $L$  can be moved by integration by parts from one side of the integral to the other,

$$- \int_a^b X^* L[X] dx = - \int_a^b (L[X])^* X dx,$$

then all the allowed *eigenvalues*  $\lambda$  must be real. (Here it is understood that  $X(x)$  satisfies the boundary conditions of the problem, though not necessarily the differential equation. An operator with this integration-by-parts symmetry is called

*self-adjoint.*) If, in addition, an intermediate step in the integration by parts is a manifestly positive (or nonnegative) integral, then the  $\lambda$ 's must be positive (or nonnegative, respectively).

To summarize, in the one-dimensional heat problem with Dirichlet boundary conditions we have found the eigenvalues

$$\lambda_n \equiv \omega_n^2 = (n\pi)^2$$

and the corresponding solutions

$$w_{\text{sep}}(t, x) = \sin(\omega_n x) e^{-\omega_n^2 t}.$$

We still need to investigate how to superpose such solutions to obtain a solution with the arbitrary initial data  $w(0, x) = g(x)$ . So, let us assume that such a solution exists, and see if that assumption leads us either to useful information (good), or to a contradiction (bad):

$$w(t, x) = \sum_{n=1}^{\infty} b_n \sin(\omega_n x) e^{-\omega_n^2 t}$$

for some (not yet known) coefficients  $b_n$ . Then

$$g(x) = \sum_{n=1}^{\infty} b_n \sin(\omega_n x). \quad (*)$$

This is supposed to hold on the interval  $0 < x < 1$ .

More generally, if the spatial interval is  $0 < x < L$ , then we would like (\*) to be true for the appropriate choice of the  $\omega_n$ 's — namely,

$$\omega_n = \frac{n\pi}{L}$$

(the positive solutions of  $0 = X(L) = \sin(\omega L)$ ). In particular, if  $L = \pi$ , then  $\omega_n = n$ . I shall develop the theory of (\*) for the case  $L = \pi$ , rather than the case  $L = 1$  that I've been discussing heretofore. (It makes the formulas simpler.)

To find the  $b_n$  in (\*) we multiply that equation by  $\sin mx$  and integrate from 0 to  $\pi$ . We assume that the integral of the infinite series exists and is equal to the sum of the integrals of the individual terms:

$$\int_0^{\pi} g(x) \sin mx \, dx = \sum_{n=1}^{\infty} b_n \int_0^{\pi} \sin nx \sin mx \, dx.$$

(In the general case, of course, the integral would be from 0 to  $L$ .) Now

$$\sin nx \sin mx = \frac{1}{2} \cos (nx - mx) - \frac{1}{2} \cos (nx + mx),$$

so

$$\int_0^\pi \sin nx \sin mx dx = \left[ \frac{1}{2(n-m)} \sin (n-m)x - \frac{1}{2(n+m)} \sin (n+m)x \right]_0^\pi = 0$$

— provided that  $n \neq m$ . If  $n = m$  we have

$$\int_0^\pi \sin^2 mx dx = \left[ \frac{1}{2} x - \frac{1}{4m} \sin (2mx) \right]_0^\pi = \frac{\pi}{2}.$$

Thus only the  $n = m$  term in the sum survives, and

$$\int_0^\pi g(x) \sin mx dx = \frac{\pi}{2} b_m.$$

Conclusion: If (\*) is true,

$$g(x) = \sum_{n=1}^{\infty} b_n \sin nx,$$

then

$$b_n = \frac{2}{\pi} \int_0^\pi g(x) \sin nx dx.$$

(\*) is called the *Fourier sine series* of the function  $g$ , and the  $b_n$  are its *Fourier coefficients*.