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):

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}\,, \qquad \text{(PDE)}$$

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

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

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

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

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

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

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

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 $-K$:

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

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

$$X(0) = 0 = X(L). \tag{$*$}$$

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,

$$X'' + KX = 0, \tag{$\dagger$}$$

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 $x = 0$ gives

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

We see that we can get a nontrivial solution if λL is a place where the sine function equals zero. Well, $\sin z = 0$ if and only if $z = 0$, π, 2π, $\ldots$, or $-\pi$, -2π, $\ldots$. 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 ns 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

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

The Xs and λs 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?

$$f(x) = u(x,0) = X(x)T(0) = C\sin(\lambda x),$$

$$g(x) = \frac{\partial u}{\partial t}(x,0) = X(x)T'(0) = c\lambda D\sin(\lambda x).$$

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

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

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 a 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

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

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

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

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

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

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 ($u_{\rm sep} = X(x)T(t)$) 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.