1

WHERE PDEs COME FROM

After thinking about the meaning of a partial differential equation, we will flex our mathematical muscles by solving a few of them. Then we will see how naturally they arise in the physical sciences. The physics will motivate the formulation of boundary conditions and initial conditions.

1.1 WHAT IS A PARTIAL DIFFERENTIAL EQUATION?

The key defining property of a partial differential equation (PDE) is that there is more than one independent variable $x, y, \ldots$. There is a dependent variable that is an unknown function of these variables $u(x, y, \ldots)$. We will often denote its derivatives by subscripts; thus $\partial u/\partial x = u_x$, and so on. A PDE is an identity that relates the independent variables, the dependent variable u, and the partial derivatives of u. It can be written as

$$F(x, y, u(x, y), u_x(x, y), u_y(x, y)) = F(x, y, u, u_x, u_y) = 0. \qquad (1)$$

This is the most general PDE in two independent variables of *first* order. The *order* of an equation is the highest derivative that appears. The most general *second*-order PDE in two independent variables is

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0. \qquad (2)$$

A *solution* of a PDE is a function $u(x, y, \ldots)$ that satisfies the equation identically, at least in some region of the $x, y, \ldots$ variables.

When solving an ordinary differential equation (ODE), one sometimes reverses the roles of the independent and the dependent variables—for instance, for the separable ODE $\frac{du}{dx} = u^3$. For PDEs, the distinction between the independent variables and the dependent variable (the unknown) is always maintained.

Some examples of PDEs (all of which occur in physical theory) are:

1. $u_x + u_y = 0$ (transport)
2. $u_x + yu_y = 0$ (transport)
3. $u_x + uu_y = 0$ (shock wave)
4. $u_{xx} + u_{yy} = 0$ (Laplace's equation)
5. $u_{tt} - u_{xx} + u^3 = 0$ (wave with interaction)
6. $u_t + uu_x + u_{xxx} = 0$ (dispersive wave)
7. $u_{tt} + u_{xxxx} = 0$ (vibrating bar)
8. $u_t - iu_{xx} = 0$ $(i = \sqrt{-1})$ (quantum mechanics)

Each of these has two independent variables, written either as x and y or as x and t. Examples 1 to 3 have order one; 4, 5, and 8 have order two; 6 has order three; and 7 has order four. Examples 3, 5, and 6 are distinguished from the others in that they are not "linear." We shall now explain this concept.

Linearity means the following. Write the equation in the form $\mathscr{L}u = 0$, where $\mathscr{L}$ is an *operator*. That is, if v is any function, $\mathscr{L}v$ is a new function. For instance, $\mathscr{L} = \partial/\partial x$ is the operator that takes v into its partial derivative v_x. In Example 2, the operator $\mathscr{L}$ is $\mathscr{L} = \partial/\partial x + y\partial/\partial y$. ($\mathscr{L}u = u_x + yu_y$.) The definition we want for linearity is

$$\mathscr{L}(u + v) = \mathscr{L}u + \mathscr{L}v \qquad \mathscr{L}(cu) = c\mathscr{L}u \tag{3}$$

for any functions u, v and any constant c. Whenever (3) holds (for all choices of u, v, and c), $\mathscr{L}$ is called *linear operator*. The equation

$$\mathscr{L}u = 0 \tag{4}$$

is called *linear* if $\mathscr{L}$ is a linear operator. Equation (4) is called a *homogeneous linear equation*. The equation

$$\mathscr{L}u = g, \tag{5}$$

where $g \neq 0$ is a given function of the independent variables, is called an *inhomogeneous linear equation*. For instance, the equation

$$(\cos xy^2)u_x - y^2 u_y = \tan(x^2 + y^2) \tag{6}$$

is an inhomogeneous linear equation.

As you can easily verify, five of the eight equations above are linear as well as homogeneous. Example 5, on the other hand, is not linear because although $(u + v)_{xx} = u_{xx} + v_{xx}$ and $(u + v)_{tt} = u_{tt} + v_{tt}$ satisfy property (3), the cubic term does not:

$$(u + v)^3 = u^3 + 3u^2v + 3uv^2 + v^3 \neq u^3 + v^3.$$

The advantage of linearity for the equation $\mathscr{L}u = 0$ is that if u and v are both solutions, so is $(u + v)$. If $u_1, \ldots, u_n$ are all solutions, so is any linear combination

$$c_1u_1(x) + \cdots + c_nu_n(x) = \sum_{j=1}^{n} c_ju_j(x) \qquad (c_j = \text{ constants}).$$

(This is sometimes called the superposition principle.) Another consequence of linearity is that if you add a homogeneous solution [a solution of (4)] to an inhomogeneous solution [a solution of (5)], you get an inhomogeneous solution. (Why?) The mathematical structure that deals with linear combinations and linear operators is the vector space. Exercises 5–10 are review problems on vector spaces.

We'll study, almost exclusively, linear systems with constant coefficients. Recall that for ODEs you get linear combinations. The coefficients are the arbitrary constants. *For an ODE of order m, you get m arbitrary constants.*

Let's look at some PDEs.

Example 1.

Find all $u(x, y)$ satisfying the equation $u_{xx} = 0$. Well, we can integrate once to get $u_x =$ constant. But that's not really right since there's another variable y. What we really get is $u_x(x, y) = f(y)$, where $f(y)$ is arbitrary. Do it again to get $u(x, y) = f(y)x + g(y)$. This is the solution formula. Note that *there are two arbitrary* functions *in the solution*. We see this as well in the next two examples. □

Example 2.

Solve the PDE $u_{xx} + u = 0$. Again, it's really an ODE with an extra variable y. We know how to solve the ODE, so the solution is

$$u = f(y)\cos x + g(y)\sin x,$$

where again $f(y)$ and $g(y)$ are two arbitrary functions of y. You can easily check this formula by differentiating twice to verify that $u_{xx} = -u$. □

Example 3.

Solve the PDE $u_{xy} = 0$. This isn't too hard either. First let's integrate in x, regarding y as fixed. So we get

$$u_y(x, y) = f(y).$$

Next let's integrate in y regarding x as fixed. We get the solution

$$u(x, y) = F(y) + G(x),$$

where $F' = f$. □

Moral *A PDE has arbitrary functions in its solution.* In these examples the arbitrary functions are functions of one variable that combine to produce a function $u(x, y)$ of two variables which is only partly arbitrary.

A function of two variables contains *immensely* more information than a function of only one variable. Geometrically, it is obvious that a surface $\{u = f(x, y)\}$, the graph of a function of two variables, is a much more complicated object than a curve $\{u = f(x)\}$, the graph of a function of one variable.

To illustrate this, we can ask how a computer would record a function $u = f(x)$. Suppose that we choose 100 points to describe it using equally spaced values of x: $x_1, x_2, x_3, \ldots, x_{100}$. We could write them down in a column, and next to each x_j we could write the corresponding value $u_j = f(x_j)$. Now how about a function $u = f(x, y)$? Suppose that we choose 100 equally spaced values of x and also of y: $x_1, x_2, x_3, \ldots, x_{100}$ and $y_1, y_2, y_3, \ldots, y_{100}$. Each *pair* x_i, y_j provides a value $u_{ij} = f(x_i, y_j)$, so there will be $100^2 = 10{,}000$ lines of the form

$$x_i \qquad y_j \qquad u_{ij}$$

required to describe the function! (If we had a prearranged system, we would need to record only the values u_{ij}.) A function of three variables described discretely by 100 values in each variable would require a million numbers!

To understand this book what do you have to know from calculus? Certainly all the basic facts about partial derivatives and multiple integrals. For a brief discussion of such topics, see the Appendix. Here are a few things to keep in mind, some of which may be new to you.

1. Derivatives are *local*. For instance, to calculate the derivative $(\partial u/\partial x)(x_0, t_0)$ at a particular point, you need to know just the values of $u(x, t_0)$ for x *near* x_0, since the derivative is the limit as $x \to x_0$.
2. Mixed derivatives are equal: $u_{xy} = u_{yx}$. (We assume throughout this book, unless stated otherwise, that all derivatives exist and are continuous.)
3. The chain rule is used frequently in PDEs; for instance,
$$\frac{\partial}{\partial x}[f(g(x, t))] = f'(g(x, t)) \cdot \frac{\partial g}{\partial x}(x, t).$$
4. For the integrals of derivatives, the reader should learn or review Green's theorem and the divergence theorem. (See the end of Section A.3 in the Appendix.)
5. Derivatives of integrals like $I(t) = \int_{a(t)}^{b(t)} f(x, t)\, dx$ (see Section A.3).
6. Jacobians (change of variable in a double integral) (see Section A.1).
7. Infinite series of functions and their differentiation (see Section A.2).
8. Directional derivatives (see Section A.1).
9. We'll often reduce PDEs to ODEs, so we must know how to solve simple ODEs. But we won't need to know anything about tricky ODEs.

EXERCISES

1. Verify the linearity and nonlinearity of the eight examples of PDEs given in the text, by checking whether or not equations (3) are valid.
2. Which of the following operators are linear?
   (a) $\mathscr{L}u = u_x + xu_y$
   (b) $\mathscr{L}u = u_x + uu_y$
   (c) $\mathscr{L}u = u_x + u_y^2$
   (d) $\mathscr{L}u = u_x + u_y + 1$
   (e) $\mathscr{L}u = \sqrt{1+x^2}\,(\cos y)u_x + u_{yxy} - [\arctan(x/y)]u$
3. For each of the following equations, state the order and whether it is nonlinear, linear inhomogeneous, or linear homogeneous; provide reasons.
   (a) $u_t - u_{xx} + 1 = 0$
   (b) $u_t - u_{xx} + xu = 0$
   (c) $u_t - u_{xxt} + uu_x = 0$
   (d) $u_{tt} - u_{xx} + x^2 = 0$
   (e) $iu_t - u_{xx} + u/x = 0$
   (f) $u_x(1+u_x^2)^{-1/2} + u_y(1+u_y^2)^{-1/2} = 0$
   (g) $u_x + e^y u_y = 0$
   (h) $u_t + u_{xxxx} + \sqrt{1+u} = 0$
4. Show that the difference of two solutions of an inhomogeneous linear equation $\mathscr{L}u = g$ with the same g is a solution of the homogeneous equation $\mathscr{L}u = 0$.
5. Which of the following collections of 3-vectors $[a, b, c]$ are vector spaces? Provide reasons.
   (a) The vectors with $b = 0$.
   (b) The vectors with $b = 1$.
   (c) The vectors with $ab = 0$.
   (d) All the linear combinations of the two vectors $[1, 1, 0]$ and $[2, 0, 1]$.
   (e) All the vectors such that $c - a = 2b$.
6. Are the three vectors $[1, 2, 3]$, $[-2, 0, 1]$, and $[1, 10, 17]$ linearly dependent or independent? Do they span all vectors or not?
7. Are the functions $1 + x$, $1 - x$, and $1 + x + x^2$ linearly dependent or independent? Why?
8. Find a vector that, together with the vectors $[1, 1, 1]$ and $[1, 2, 1]$, forms a basis of $\mathbb{R}^3$.
9. Show that the functions $(c_1 + c_2 \sin^2 x + c_3 \cos^2 x)$ form a vector space. Find a basis of it. What is its dimension?
10. Show that the solutions of the differential equation $u''' - 3u'' + 4u = 0$ form a vector space. Find a basis of it.
11. Verify that $u(x, y) = f(x)g(y)$ is a solution of the PDE $uu_{xy} = u_x u_y$ for all pairs of (differentiable) functions f and g of one variable.

12. Verify by direct substitution that

$$u_n(x, y) = \sin nx \sinh ny$$

is a solution of $u_{xx} + u_{yy} = 0$ for every $n > 0$.

1.2 FIRST-ORDER LINEAR EQUATIONS

We begin our study of PDEs by solving some simple ones. The solution is quite geometric in spirit.

The simplest possible PDE is $\partial u/\partial x = 0$ [where $u = u(x, y)$]. Its general solution is $u = f(y)$, where f is any function of *one* variable. For instance, $u = y^2 - y$ and $u = e^y \cos y$ are two solutions. Because the solutions don't depend on x, they are constant on the lines $y =$ constant in the xy plane.

THE CONSTANT COEFFICIENT EQUATION

Let us solve

$$\boxed{au_x + bu_y = 0,} \tag{1}$$

where a and b are constants not both zero.

Geometric Method The quantity $au_x + bu_y$ is the directional derivative of u in the direction of the vector $\mathbf{V} = (a, b) = a\mathbf{i} + b\mathbf{j}$. It must always be zero. This means that $u(x, y)$ must be constant in the direction of $\mathbf{V}$. The vector $(b, -a)$ is orthogonal to $\mathbf{V}$. The lines parallel to $\mathbf{V}$ (see Figure 1) have the equations $bx - ay =$ constant. (They are called the *characteristic lines*.) The solution is constant on each such line. Therefore, $u(x, y)$ depends on $bx - ay$ only. Thus the solution is

$$\boxed{u(x, y) = f(bx - ay),} \tag{2}$$

where f is any function of one variable. Let's explain this conclusion more explicitly. On the line $bx - ay = c$, the solution u has a constant value. Call

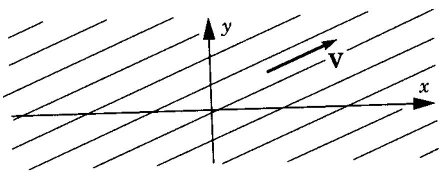

Figure 1

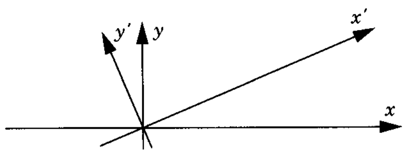

Figure 2

this value $f(c)$. Then $u(x, y) = f(c) = f(bx - ay)$. Since c is arbitrary, we have formula (2) for all values of x and y. In xyu space the solution defines a surface that is made up of parallel horizontal straight lines like a sheet of corrugated iron.

Coordinate Method Change variables (or "make a change of coordinates"; Figure 2) to

$$x' = ax + by \qquad y' = bx - ay. \tag{3}$$

Replace all x and y derivatives by x' and y' derivatives. By the chain rule,

$$u_x = \frac{\partial u}{\partial x} = \frac{\partial u}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial u}{\partial y'}\frac{\partial y'}{\partial x} = au_{x'} + bu_{y'}$$

and

$$u_y = \frac{\partial u}{\partial y} = \frac{\partial u}{\partial y'}\frac{\partial y'}{\partial y} + \frac{\partial u}{\partial x'}\frac{\partial x'}{\partial y} = bu_{x'} - au_{y'}.$$

Hence $au_x + bu_y = a(au_{x'} + bu_{y'}) + b(bu_{x'} - au_{y'}) = (a^2 + b^2)u_{x'}$. So, since $a^2 + b^2 \neq 0$, the equation takes the form $u_{x'} = 0$ in the new (primed) variables. Thus the solution is $u = f(y') = f(bx - ay)$, with f an arbitrary function of *one* variable. This is exactly the same answer as before!

Example 1.

Solve the PDE $4u_x - 3u_y = 0$, together with the auxiliary condition that $u(0, y) = y^3$. By (2) we have $u(x, y) = f(-3x - 4y)$. This is the general solution of the PDE. Setting $x = 0$ yields the equation $y^3 = f(-4y)$. Letting $w = -4y$ yields $f(w) = -w^3/64$. Therefore, $u(x, y) = (3x + 4y)^3/64$.

Solutions can usually be checked much easier than they can be derived. We check this solution by simple differentiation: $u_x = 9(3x + 4y)^2/64$ and $u_y = 12(3x + 4y)^2/64$ so that $4u_x - 3u_y = 0$. Furthermore, $u(0, y) = (3 \cdot 0 + 4y)^3/64 = y^3$. $\square$

THE VARIABLE COEFFICIENT EQUATION

The equation

$$\boxed{u_x + yu_y = 0} \tag{4}$$

is linear and homogeneous but has a variable coefficient (y). We shall illustrate for equation (4) how to use the geometric method somewhat like Example 1.

The PDE (4) itself asserts that *the directional derivative in the direction of the vector* $(1, y)$ *is zero*. The curves in the xy plane with $(1, y)$ as tangent vectors have slopes y (see Figure 3). Their equations are

$$\frac{dy}{dx} = \frac{y}{1} \tag{5}$$

This ODE has the solutions

$$y = Ce^x. \tag{6}$$

These curves are called the *characteristic curves* of the PDE (4). As C is changed, the curves fill out the xy plane perfectly without intersecting. On each of the curves $u(x, y)$ is a constant because

$$\frac{d}{dx}u(x, Ce^x) = \frac{\partial u}{\partial x} + Ce^x\frac{\partial u}{\partial y} = u_x + yu_y = 0.$$

Thus $u(x, Ce^x) = u(0, Ce^0) = u(0, C)$ is independent of x. Putting $y = Ce^x$ and $C = e^{-x}y$, we have

$$u(x, y) = u(0,\ e^{-x}y).$$

It follows that

$$\boxed{u(x, y) = f(e^{-x}y)} \tag{7}$$

is the *general solution* of this PDE, where again f is an arbitrary function of only a single variable. This is easily checked by differentiation using the chain rule (see Exercise 4). Geometrically, the "picture" of the solution $u(x, y)$ is that it is *constant on each characteristic curve* in Figure 3.

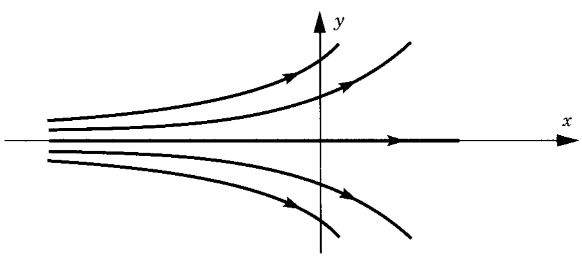

Figure 3

Example 2.

Find the solution of (4) that satisfies the auxiliary condition $u(0, y) = y^3$. Indeed, putting $x = 0$ in (7), we get $y^3 = f(e^{-0}y)$, so that $f(y) = y^3$. Therefore, $u(x, y) = (e^{-x}y)^3 = e^{-3x}y^3$. □

Example 3.

Solve the PDE

$$\boxed{u_x + 2xy^2u_y = 0.} \tag{8}$$

The characteristic curves satisfy the ODE $dy/dx = 2xy^2/1 = 2xy^2$. To solve the ODE, we separate variables: $dy/y^2 = 2x\,dx$; hence $-1/y = x^2 - C$, so that

$$y = (C - x^2)^{-1}. \tag{9}$$

These curves are the characteristics. Again, $u(x, y)$ is a constant on each such curve. (Check it by writing it out.) So $u(x, y) = f(C)$, where f is an arbitrary function. Therefore, the general solution of (8) is obtained by solving (9) for C. That is,

$$\boxed{u(x, y) = f\left(x^2 + \frac{1}{y}\right).} \tag{10}$$

Again this is easily checked by differentiation, using the chain rule: $u_x = 2x \cdot f'(x^2 + 1/y)$ and $u_y = -(1/y^2) \cdot f'(x^2 + 1/y)$, whence $u_x + 2xy^2u_y = 0$. □

In summary, the geometric method works nicely for any PDE of the form $a(x, y)u_x + b(x, y)u_y = 0$. It reduces the solution of the PDE to the solution of the ODE $dy/dx = b(x, y)/a(x, y)$. If the ODE can be solved, so can the PDE. Every solution of the PDE is constant on the solution curves of the ODE.

Moral Solutions of PDEs generally depend on arbitrary functions (instead of arbitrary constants). You need an auxiliary condition if you want to determine a unique solution. Such conditions are usually called *initial* or *boundary* conditions. We shall encounter these conditions throughout the book.

EXERCISES

1. Solve the first-order equation $2u_t + 3u_x = 0$ with the auxiliary condition $u = \sin x$ when $t = 0$.
2. Solve the equation $3u_y + u_{xy} = 0$. (*Hint*: Let $v = u_y$.)

3. Solve the equation $(1 + x^2)u_x + u_y = 0$. Sketch some of the characteristic curves.
4. Check that (7) indeed solves (4).
5. Solve the equation $xu_x + yu_y = 0$.
6. Solve the equation $\sqrt{1 - x^2}u_x + u_y = 0$ with the condition $u(0, y) = y$.
7. (a) Solve the equation $yu_x + xu_y = 0$ with $u(0, y) = e^{-y^2}$.
   (b) In which region of the xy plane is the solution uniquely determined?
8. Solve $au_x + bu_y + cu = 0$.
9. Solve the equation $u_x + u_y = 1$.
10. Solve $u_x + u_y + u = e^{x+2y}$ with $u(x, 0) = 0$.
11. Solve $au_x + bu_y = f(x, y)$, where $f(x, y)$ is a given function. If $a \neq 0$, write the solution in the form
$$u(x, y) = (a^2 + b^2)^{-1/2} \int_L f\, ds + g(bx - ay),$$
where g is an arbitrary function of one variable, L is the characteristic line segment from the y axis to the point (x, y), and the integral is a line integral. (*Hint:* Use the coordinate method.)
12. Show that the new coordinate axes defined by (3) are orthogonal.
13. Use the coordinate method to solve the equation
$$u_x + 2u_y + (2x - y)u = 2x^2 + 3xy - 2y^2.$$

1.3 FLOWS, VIBRATIONS, AND DIFFUSIONS

The subject of PDEs was practically a branch of physics until the twentieth century. In this section we present a series of examples of PDEs as they occur in physics. They provide the basic motivation for all the PDE problems we study in the rest of the book. We shall see that most often in physical problems the independent variables are those of space x, y, z, and time t.

Example 1. Simple Transport

Consider a fluid, water, say, flowing at a constant rate c along a horizontal pipe of fixed cross section in the positive x direction. A substance, say a pollutant, is suspended in the water. Let $u(x, t)$ be its concentration in grams/centimeter at time t. Then

$$\boxed{u_t + cu_x = 0.} \tag{1}$$

(That is, the rate of change u_t of concentration is proportional to the gradient u_x. Diffusion is assumed to be negligible.) Solving this equation as in Section 1.2, we find that the concentration is a function of $(x - ct)$

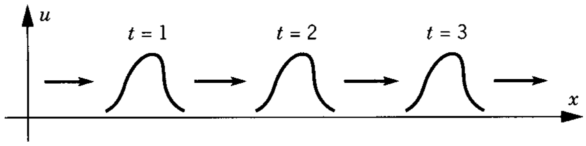

Figure 1

only. This means that the substance is transported to the right at a fixed speed c. Each individual particle moves to the right at speed c; that is, in the xt plane, it moves precisely along a characteristic line (see Figure 1). □

Derivation of Equation (1). The amount of pollutant in the interval $[0, b]$ at the time t is $M = \int_0^b u(x,t)\,dx$, in grams, say. At the later time $t + h$, the same molecules of pollutant have moved to the right by $c \cdot h$ centimeters. Hence

$$M = \int_0^b u(x,t)dx = \int_{ch}^{b+ch} u(x, t+h)\,dx.$$

Differentiating with respect to b, we get

$$u(b,t) = u(b + ch, t + h).$$

Differentiating with respect to h and putting $h = 0$, we get

$$0 = cu_x(b,t) + u_t(b,t),$$

which is equation (1). □

Example 2. Vibrating String

Consider a flexible, elastic homogenous string or thread of length l, which undergoes relatively small transverse vibrations. For instance, it could be a guitar string or a plucked violin string. At a given instant t, the string might look as shown in Figure 2. Assume that it remains in a plane. Let $u(x, t)$ be its displacement from equilibrium position at time t and position x. Because the string is perfectly flexible, the tension (force) is directed tangentially along the string (Figure 3). Let $T(x, t)$ be the magnitude of this tension vector. Let ρ be the density (mass per unit length) of the string. It is a constant because the string is homogeneous. We shall write down Newton's law for the part of the string between any two points at $x = x_0$ and $x = x_1$. The slope of the string at x_1 is

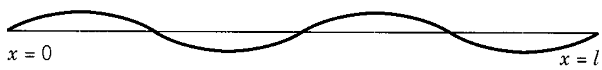

Figure 2

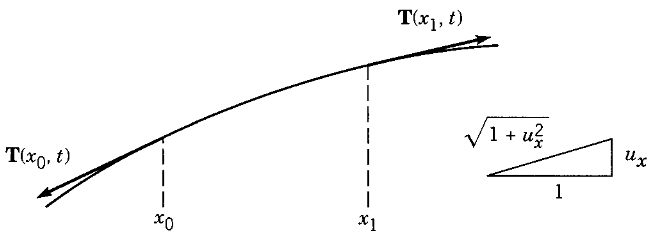

Figure 3

$u_x(x_1, t)$. Newton's law $\mathbf{F} = m\mathbf{a}$ in its longitudinal (x) and transverse (u) components is

$$\left.\frac{T}{\sqrt{1+u_x^2}}\right|_{x_0}^{x_1} = 0 \qquad \text{(longitudinal)}$$

$$\left.\frac{Tu_x}{\sqrt{1+u_x^2}}\right|_{x_0}^{x_1} = \int_{x_0}^{x_1} \rho u_{tt}\,dx \qquad \text{(transverse)}$$

The right sides are the components of the mass times the acceleration integrated over the piece of string. Since we have assumed that the motion is purely transverse, there is no longitudinal motion.

Now we also assume that the motion is small—more specifically, that $|u_x|$ is quite small. Then $\sqrt{1+u_x^2}$ may be replaced by 1. This is justified by the Taylor expansion, actually the binomial expansion,

$$\sqrt{1+u_x^2} = 1 + \tfrac{1}{2}u_x^2 + \cdots$$

where the dots represent higher powers of u_x. If u_x is small, it makes sense to drop the even smaller quantity u_x^2 and its higher powers. With the square roots replaced by 1, the first equation then says that T is constant along the string. Let us assume that T is independent of t as well as x. The second equation, differentiated, says that

$$(Tu_x)_x = \rho u_{tt}.$$

That is,

$$\boxed{u_{tt} = c^2 u_{xx} \quad \text{where } c = \sqrt{\frac{T}{\rho}}.} \qquad (2)$$

This is the *wave equation*. At this point it is not clear why c is defined in this manner, but shortly we'll see that c is the *wave speed*. □

There are many *variations* of this equation:

(i) If significant air resistance r is present, we have an extra term proportional to the speed u_t, thus:

$$u_{tt} - c^2 u_{xx} + r u_t = 0 \quad \text{where } r > 0. \tag{3}$$

(ii) If there is a transverse elastic force, we have an extra term proportional to the displacement u, as in a coiled spring, thus:

$$u_{tt} - c^2 u_{xx} + ku = 0 \quad \text{where } k > 0. \tag{4}$$

(iii) If there is an externally applied force, it appears as an extra term, thus:

$$u_{tt} - c^2 u_{xx} = f(x, t), \tag{5}$$

which makes the equation inhomogeneous.

Our derivation of the wave equation has been quick but not too precise. A much more careful derivation can be made, which makes precise the physical and mathematical assumptions [We, Chap. 1].

The same wave equation or a variation of it describes many other wavelike phenomena, such as the vibrations of an elastic bar, the sound waves in a pipe, and the long water waves in a straight canal. Another example is the equation for the electrical current in a transmission line,

$$u_{xx} = CLu_{tt} + (CR + GL)u_t + GRu,$$

where C is the capacitance per unit length, G the leakage resistance per unit length, R the resistance per unit length, and L the self-inductance per unit length.

Example 3. Vibrating Drumhead

The two-dimensional version of a string is an elastic, flexible, homogeneous drumhead, that is, a membrane stretched over a frame. Say the frame lies in the xy plane (see Figure 4), $u(x, y, t)$ is the vertical

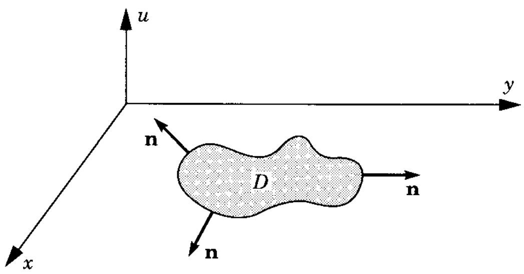

Figure 4

displacement, and there is no horizontal motion. The horizontal components of Newton's law again give constant tension T. Let D be any domain in the xy plane, say a circle or a rectangle. Let bdy D be its boundary curve. We use reasoning similar to the one-dimensional case. The vertical component gives (approximately)

$$F = \int_{\text{bdy } D} T \frac{\partial u}{\partial n}\, ds = \iint_D \rho u_{tt}\, dx\, dy = ma,$$

where the left side is the total force acting on the piece D of the membrane, and where $\partial u/\partial n = \mathbf{n} \cdot \nabla u$ is the directional derivative in the outward normal direction, $\mathbf{n}$ being the unit outward normal vector on bdy D. By Green's theorem (see Section A.3 in the Appendix), this can be rewritten as

$$\iint_D \nabla \cdot (T \nabla u)\, dx\, dy = \iint_D \rho u_{tt}\, dx\, dy.$$

Since D is arbitrary, we deduce from the second vanishing theorem in Section A.1 that $\rho u_{tt} = \nabla \cdot (T \nabla u)$. Since T is constant, we get

$$\boxed{u_{tt} = c^2 \nabla \cdot (\nabla u) \equiv c^2 (u_{xx} + u_{yy}),} \tag{6}$$

where $c = \sqrt{T/\rho}$ as before and $\nabla \cdot (\nabla u) = \operatorname{div} \operatorname{grad} u = u_{xx} + u_{yy}$ is known as the *two-dimensional laplacian*. Equation (6) is the two-dimensional wave equation. □

The pattern is now clear. Simple three-dimensional vibrations obey the equation

$$\boxed{u_{tt} = c^2 (u_{xx} + u_{yy} + u_{zz}).} \tag{7}$$

The operator $\mathscr{L} = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial/\partial z^2$ is called the *three-dimensional laplacian* operator, usually denoted by Δ or ∇^2. Physical examples described by the three-dimensional wave equation or a variation of it include the vibrations of an elastic solid, sound waves in air, electromagnetic waves (light, radar, etc.), linearized supersonic airflow, free mesons in nuclear physics, and seismic waves propagating through the earth.

Example 4. Diffusion

Let us imagine a motionless liquid filling a straight tube or pipe and a chemical substance, say a dye, which is diffusing through the liquid. Simple diffusion is characterized by the following law. [It is not to

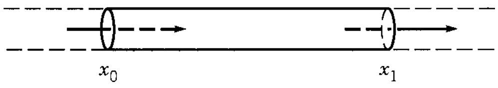

Figure 5

be confused with convection (transport), which refers to currents in the liquid.] The dye moves from regions of higher concentration to regions of lower concentration. The rate of motion is proportional to the concentration gradient. (This is known as Fick's law of diffusion.) Let $u(x, t)$ be the concentration (mass per unit length) of the dye at position x of the pipe at time t.

In the section of pipe from x_0 to x_1 (see Figure 5), the mass of dye is

$$M(t) = \int_{x_0}^{x_1} u(x, t)\, dx, \quad \text{so} \quad \frac{dM}{dt} = \int_{x_0}^{x_1} u_t(x, t)\, dx.$$

The mass in this section of pipe cannot change except by flowing in or out of its ends. By Fick's law,

$$\frac{dM}{dt} = \text{flow in} - \text{flow out} = ku_x(x_1, t) - ku_x(x_0, t),$$

where k is a proportionality constant. Therefore, those two expressions are equal:

$$\int_{x_0}^{x_1} u_t(x, t)\, dx = ku_x(x_1, t) - ku_x(x_0, t).$$

Differentiating with respect to x_1, we get

$$\boxed{u_t = ku_{xx}.} \tag{8}$$

This is the *diffusion equation*.

In three dimensions we have

$$\iiint_D u_t\, dx\, dy\, dz = \iint_{\text{bdy } D} k(\mathbf{n} \cdot \nabla u)\, dS,$$

where D is any solid domain and bdy D is its bounding surface. By the divergence theorem (using the arbitrariness of D as in Example 3), we get the *three-dimensional diffusion equation*

$$\boxed{u_t = k(u_{xx} + u_{yy} + u_{zz}) = k\, \Delta u.} \tag{9}$$

If there is an external source (or a "sink") of the dye, and if the rate k of diffusion is a variable, we get the more general inhomogeneous equation

$$u_t = \nabla \cdot (k\, \nabla u) + f(x,\ t).$$

The same equation describes the conduction of heat, brownian motion, diffusion models of population dynamics, and many other phenomena. □

Example 5. Heat Flow

We let $u(x, y, z, t)$ be the temperature and let $H(t)$ be the amount of heat (in calories, say) contained in a region D. Then

$$H(t) = \iiint\limits_D c\rho u \, dx\, dy\, dz,$$

where c is the "specific heat" of the material and ρ is its density (mass per unit volume). The change in heat is

$$\frac{dH}{dt} = \iiint\limits_D c\rho u_t \, dx\, dy\, dz.$$

Fourier's law says that heat flows from hot to cold regions proportionately to the temperature gradient. But heat cannot be lost from D except by leaving it through the boundary. This is the law of conservation of energy. Therefore, the change of heat energy in D also equals the heat flux across the boundary,

$$\frac{dH}{dt} = \iint\limits_{\text{bdy } D} \kappa(\mathbf{n} \cdot \nabla u)\, dS,$$

where κ is a proportionality factor (the "heat conductivity"). By the divergence theorem,

$$\iiint\limits_D c\rho \frac{\partial u}{\partial t}\, dx\, dy\, dz = \iiint\limits_D \nabla \cdot (\kappa\, \nabla u)\, dx\, dy\, dz$$

and we get the *heat equation*

$$\boxed{c\rho \frac{\partial u}{\partial t} = \nabla \cdot (\kappa\, \nabla u).} \tag{10}$$

If c, ρ, and κ are constants, it is exactly the same as the diffusion equation! □

Example 6. Stationary Waves and Diffusions

Consider any of the four preceding examples in a situation where the physical state does not change with time. Then $u_t = u_{tt} = 0$. So *both* the wave *and* the diffusion equations reduce to

$$\boxed{\Delta u = u_{xx} + u_{yy} + u_{zz} = 0.} \tag{11}$$

This is called the *Laplace equation*. Its solutions are called *harmonic functions*. For example, consider a hot object that is constantly heated in an oven. The heat is not expected to be evenly distributed throughout the oven. The temperature of the object eventually reaches a steady (or equilibrium) state. This is a harmonic function $u(x, y, z)$. (Of course, if the heat were being supplied evenly in all directions, the steady state would be $u \equiv$ constant.) In the *one*-dimensional case (e.g., a laterally insulated thin rod that exchanges heat with its environment only through its ends), we would have u a function of x only. So the Laplace equation would reduce simply to $u_{xx} = 0$. Hence $u = c_1 x + c_2$. The two- and three-dimensional cases are *much* more interesting (see Chapter 6 for the solutions). □

Example 7. The Hydrogen Atom

This is an electron moving around a proton. Let m be the mass of the electron, e its charge, and h Planck's constant divided by 2π. Let the origin of coordinates (x, y, z) be at the proton and let $r = (x^2 + y^2 + z^2)^{1/2}$ be the spherical coordinate. Then the motion of the electron is given by a "wave function" $u(x, y, z, t)$ which satisfies Schrödinger's equation

$$\boxed{-ihu_t = \frac{h^2}{2m}\Delta u + \frac{e^2}{r}u} \tag{12}$$

in all of space $-\infty < x,y,z < +\infty$. Furthermore, we are supposed to have $\iiint |u|^2 dx\, dy\, dz = 1$ (integral over all space). Note that $i = \sqrt{-1}$ and u is complex-valued. The coefficient function e^2/r is called the potential. For any other atom with a single electron, such as a helium ion, e^2 is replaced by Ze^2, where Z is the atomic number. □

What does this mean physically? In quantum mechanics quantities cannot be measured exactly but only with a certain probability. The *wave function* $u(x, y, z, t)$ represents a possible *state* of the electron. If D is *any* region in xyz space, then

$$\iiint_D |u|^2\, dx\, dy\, dz$$

is the probability of finding the electron in the region D at the time t. The *expected z coordinate of the position* of the electron at the time t is the value of the integral

$$\iiint z|u(x, y, z, t)|^2 \, dx \, dy \, dz;$$

similarly for the x and y coordinates. The *expected z coordinate of the momentum* is

$$\iiint -ih\frac{\partial u}{\partial z}(x, y, z, t) \cdot \bar{u}(x, y, z, t) \, dx \, dy \, dz,$$

where $\bar{u}$ is the complex conjugate of u. All other observable quantities are given by operators A, which act on functions. The expected value of the observable A equals

$$\iiint Au(x, y, z, t) \cdot \bar{u}(x, y, z, t) \, dx \, dy \, dz.$$

Thus the position is given by the operator $Au = \mathbf{x}u$, where $\mathbf{x} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, and the momentum is given by the operator $Au = -ih\nabla u$.

Schrödinger's equation is most easily regarded simply as an axiom that leads to the correct physical conclusions, rather than as an equation that can be derived from simpler principles. It explains why atoms are stable and don't collapse. It explains the energy levels of the electron in the hydrogen atom observed by Bohr. In principle, elaborations of it explain the structure of *all* atoms and molecules and so all of chemistry! With many particles, the wave function u depends on time t and all the coordinates of all the particles and so is a function of a large number of variables. The Schrödinger equation then becomes

$$-ihu_t = \sum_{i=1}^{n} \frac{h^2}{2m_i}(u_{x_i x_i} + u_{y_i y_i} + u_{z_i z_i}) + V(x_1, \ldots, z_n)u$$

for n particles, where the potential function V depends on all the $3n$ coordinates. Except for the hydrogen and helium atoms (the latter having two electrons), the mathematical analysis is impossible to carry out completely and cannot be calculated even with the help of the modern computer. Nevertheless, with the use of various approximations, many of the facts about more complicated atoms and the chemical binding of molecules can be understood. $\square$

This has been a brief introduction to the sources of PDEs in physical problems. Many realistic situations lead to much more complicated PDEs. See Chapter 13 for some additional examples.

EXERCISES

1. Carefully derive the equation of a string in a medium in which the resistance is proportional to the velocity.
2. A flexible chain of length l is hanging from one end $x = 0$ but oscillates horizontally. Let the x axis point downward and the u axis point to the right. Assume that the force of gravity at each point of the chain equals the weight of the part of the chain below the point and is directed tangentially along the chain. Assume that the oscillations are small. Find the PDE satisfied by the chain.
3. On the sides of a thin rod, heat exchange takes place (obeying Newton's law of cooling—flux proportional to temperature difference) with a medium of constant temperature T_0. What is the equation satisfied by the temperature $u(x, t)$, neglecting its variation across the rod?
4. Suppose that some particles which are suspended in a liquid medium would be pulled down at the constant velocity $V > 0$ by gravity in the absence of diffusion. Taking account of the diffusion, find the equation for the concentration of particles. Assume homogeneity in the horizontal directions x and y. Let the z axis point upwards.
5. Derive the equation of one-dimensional diffusion in a medium that is moving along the x axis to the right at constant speed V.
6. Consider heat flow in a long circular cylinder where the temperature depends only on t and on the distance r to the axis of the cylinder. Here $r = \sqrt{x^2 + y^2}$ is the cylindrical coordinate. From the three-dimensional heat equation derive the equation $u_t = k(u_{rr} + u_r/r)$.
7. Solve Exercise 6 in a ball except that the temperature depends only on the spherical coordinate $\sqrt{x^2 + y^2 + z^2}$. Derive the equation $u_t = k(u_{rr} + 2u_r/r)$.
8. For the hydrogen atom, if $\int |u|^2 \, d\mathbf{x} = 1$ at $t = 0$, show that the same is true at all later times. (*Hint:* Differentiate the integral with respect to t, taking care about the solution being complex valued. Assume that u and $\nabla u \to 0$ fast enough as $|\mathbf{x}| \to \infty$.)
9. This is an exercise on the divergence theorem

$$\iiint_D \nabla \cdot \mathbf{F} \, d\mathbf{x} = \iint_{\text{bdy } D} \mathbf{F} \cdot \mathbf{n} \, dS,$$

valid for any bounded domain D in space with boundary surface bdy D and unit outward normal vector $\mathbf{n}$. If you never learned it, see Section A.3. It is crucial that D be bounded As an exercise, verify it in the following case by calculating both sides separately: $\mathbf{F} = r^2\mathbf{x}$, $\mathbf{x} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, $r^2 = x^2 + y^2 + z^2$, and $D =$ the ball of radius a and center at the origin.

10. If $\mathbf{f}(\mathbf{x})$ is continuous and $|\mathbf{f}(\mathbf{x})| \le 1/(|\mathbf{x}|^3 + 1)$ for all $\mathbf{x}$, show that

$$\iiint_{\text{all space}} \nabla \cdot \mathbf{f}\, d\mathbf{x} = 0.$$

(*Hint:* Take D to be a large ball, apply the divergence theorem, and let its radius tend to infinity.)

11. If curl $\mathbf{v} = \mathbf{0}$ in all of three-dimensional space, show that there exists a scalar function $\phi(x, y, z)$ such that $\mathbf{v} = \text{grad}\, \phi$.

1.4 INITIAL AND BOUNDARY CONDITIONS

Because PDEs typically have so many solutions, as we saw in Section 1.2, we single out one solution by imposing auxiliary conditions. We attempt to formulate the conditions so as to specify a unique solution. These conditions are motivated by the physics and they come in two varieties, initial conditions and boundary conditions.

An *initial condition* specifies the physical state at a particular time t_0. For the diffusion equation the initial condition is

$$u(\mathbf{x}, t_0) = \phi(\mathbf{x}), \tag{1}$$

where $\phi(\mathbf{x}) = \phi(x, y, z)$ is a given function. For a diffusing substance, $\phi(\mathbf{x})$ is the initial concentration. For heat flow, $\phi(\mathbf{x})$ is the initial temperature. For the Schrödinger equation, too, (1) is the usual initial condition.

For the wave equation there is a *pair* of initial conditions

$$u(\mathbf{x},\ t_0) = \phi(\mathbf{x}) \quad \text{and} \quad \frac{\partial u}{\partial t}(\mathbf{x},\ t_0) = \psi(\mathbf{x}), \tag{2}$$

where $\phi(\mathbf{x})$ is the initial position and $\psi(\mathbf{x})$ is the initial velocity. It is clear on physical grounds that both of them must be specified in order to determine the position $u(\mathbf{x}, t)$ at later times. (We shall also prove this mathematically.) □

In each physical problem we have seen that there is a *domain D* in which the PDE is valid. For the vibrating string, D is the interval $0 < x < l$, so the boundary of D consists only of the two points $x = 0$ and $x = l$. For the drumhead, the domain is a plane region and its boundary is a closed curve. For the diffusing chemical substance, D is the container holding the liquid, so its boundary is a surface $S = \text{bdy}\, D$. For the hydrogen atom, the domain is all of space, so it has no boundary.

It is clear, again from our physical intuition, that it is necessary to specify some *boundary condition* if the solution is to be determined. The three most important kinds of boundary conditions are:

(D) u is specified ("*Dirichlet* condition")

(N) the normal derivative $\partial u/\partial n$ is specified ("*Neumann* condition")

(R) $\partial u/\partial n + au$ is specified ("*Robin* condition")

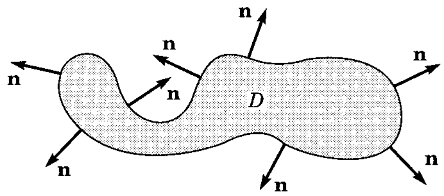

Figure 1

where a is a given function of x, y, z, and t. Each is to hold for all t and for $\mathbf{x} = (x, y, z)$ belonging to bdy D. Usually, we write (D), (N), and (R) as equations. For instance, (N) is written as the equation

$$\frac{\partial u}{\partial n} = g(\mathbf{x}, t) \tag{3}$$

where g is a given function that could be called the boundary datum. Any of these boundary conditions is called *homogeneous* if the specified function $g(\mathbf{x}, t)$ vanishes (equals zero). Otherwise, it is called inhomogenous. As usual, $\mathbf{n} = (n_1, n_2, n_3)$ denotes the unit normal vector on bdy D, which points outward from D (see Figure 1). Also, $\partial u/\partial n \equiv \mathbf{n} \cdot \nabla u$ denotes the directional derivative of u in the outward normal direction.

In *one-dimensional* problems where D is just an interval $0 < x < l$, the boundary consists of just the two endpoints, and these boundary conditions take the simple form

(D) $u(0, t) = g(t)$ and $u(l, t) = h(t)$

(N) $\dfrac{\partial u}{\partial x}(0, t) = g(t)$ and $\dfrac{\partial u}{\partial x}(l, t) = h(t)$

and similarly for the Robin condition. □

Following are some illustrations of physical problems corresponding to these boundary conditions.

THE VIBRATING STRING

If the string is held *fixed* at both ends, as for a violin string, we have the homogeneous Dirichlet conditions $u(0, t) = u(l, t) = 0$.

Imagine, on the other hand, that one end of the string is *free* to move transversally without any resistance (say, along a frictionless track); then there is no tension T at that end, so $u_x = 0$. This is a Neumann condition.

Third, the Robin condition would be the correct one if one were to imagine that an end of the string were free to move along a track but were attached to a coiled spring or rubber band (obeying Hooke's law) which tended to pull it back to equilibrium position. In that case the string would exchange some of its energy with the coiled spring.

Finally, if an end of the string were simply moved in a specified way, we would have an inhomogeneous Dirichlet condition at that end.

DIFFUSION

If the diffusing substance is enclosed in a container D so that none can escape or enter, then the concentration gradient in the normal direction must vanish, by Fick's law (see Exercise 2). Thus $\partial u/\partial n = 0$ on $S = \text{bdy } D$, which is the Neumann condition.

If, on the other hand, the container is permeable and is so constructed that any substance that escapes to the boundary of the container is immediately washed away, then we have $u = 0$ on S.

HEAT

Heat conduction is described by the diffusion equation with $u(\mathbf{x}, t) =$ temperature. If the object D through which the heat is flowing is perfectly *insulated*, then no heat flows across the boundary and we have the Neumann condition $\partial u/\partial n = 0$ (see Exercise 2).

On the other hand, if the object were immersed in a large *reservoir* of specified temperature $g(t)$ and there were perfect thermal conduction, then we'd have the Dirichlet condition $u = g(t)$ on bdy D.

Suppose that we had a uniform rod insulated along its length $0 \le x \le l$, whose end at $x = l$ were immersed in the reservoir of temperature $g(t)$. If heat were exchanged between the end and the reservoir so as to obey Newton's law of cooling, then

$$\frac{\partial u}{\partial x}(l, t) = -a[u(l, t) - g(t)],$$

where $a > 0$. Heat from the hot rod radiates into the cool reservoir. This is an inhomogeneous Robin condition.

LIGHT

Light is an electromagnetic field and as such is described by Maxwell's equations (see Chapter 13). Each component of the electric and magnetic field satisfies the wave equation. It is through the boundary conditions that the various components are related to each other. (They are "coupled.") Imagine, for example, light reflecting off a ball with a mirrored surface. This is a scattering problem. The domain D where the light is propagating is the exterior of the ball. Certain boundary conditions then are satisfied by the electromagnetic field components. When polarization effects are not being studied, some scientists use the wave equation with homogeneous Dirichlet or Neumann conditions as a considerably simplified model of such a situation.

SOUND

Our ears detect small disturbances in the air. The disturbances are described by the equations of gas dynamics, which form a system of nonlinear equations with velocity $\mathbf{v}$ and density ρ as the unknowns. But *small* disturbances are described quite well by the so-called linearized equations, which are a lot simpler; namely,

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{c_0^2}{\rho_0} \operatorname{grad} \rho = 0 \tag{4}$$

$$\frac{\partial \rho}{\partial t} + \rho_0 \operatorname{div} \mathbf{v} = 0 \tag{5}$$

(four scalar equations altogether). Here ρ_0 is the density and c_0 is the speed of sound in still air.

Assume now that the curl of $\mathbf{v}$ is zero; this means that there are no sound "eddies" and the velocity $\mathbf{v}$ is irrotational. It follows that ρ and all three components of $\mathbf{v}$ satisfy the wave equation:

$$\frac{\partial^2 \mathbf{v}}{\partial t^2} = c_0^2 \, \Delta \mathbf{v} \quad \text{and} \quad \frac{\partial^2 \rho}{\partial t^2} = c_0^2 \, \Delta \rho. \tag{6}$$

The interested reader will find a derivation of these equations in Section 13.2.

Now if we are describing sound propagation in a closed, sound-insulated room D with *rigid* walls, say a concert hall, then the air molecules at the wall can only move parallel to the boundary, so that no sound can travel in a normal direction to the boundary. So $\mathbf{v} \cdot \mathbf{n} = 0$ on bdy D. Since curl $\mathbf{v} = 0$, there is a standard fact in vector calculus (Exercise 1.3.11) which says that there is a "potential" function ψ such that $\mathbf{v} = -\operatorname{grad} \psi$. The potential also satisfies the wave equation $\partial^2\psi/\partial t^2 = c_0^2 \, \Delta\psi$, and the boundary condition for it is $-\mathbf{v} \cdot \mathbf{n} = \mathbf{n} \cdot \operatorname{grad} \psi = 0$ or Neumann's condition for ψ.

At an *open* window of the room D, the atmospheric pressure is a constant and there is no difference of pressure across the window. The pressure p is proportional to the density ρ, for small disturbances of the air. Thus ρ is a constant at the window, which means that ρ satisfies the Dirichlet boundary condition $\rho = \rho_0$.

At a *soft* wall, such as an elastic membrane covering an open window, the pressure difference $p - p_0$ across the membrane is proportional to the normal velocity $\mathbf{v} \cdot \mathbf{n}$, namely

$$p - p_0 = Z \, \mathbf{v} \cdot \mathbf{n},$$

where Z is called the acoustic impedance of the wall. (A rigid wall has a very large impedance and an open window has zero impedance.) Now $p - p_0$ is in turn proportional to $\rho - \rho_0$ for small disturbances. Thus the system of four equations (4),(5) satisfies the boundary condition

$$\mathbf{v} \cdot \mathbf{n} = a(\rho - \rho_0),$$

where a is a constant proportional to $1/Z$. (See [MI] for further discussion.) □

A different kind of boundary condition in the case of the wave equation is

$$\frac{\partial u}{\partial n} + b\frac{\partial u}{\partial t} = 0. \tag{7}$$

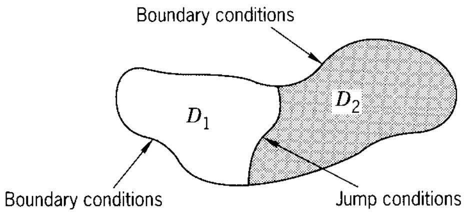

Figure 2

This condition means that energy is *radiated to* ($b > 0$) or *absorbed from* ($b < 0$) the exterior through the boundary. For instance, a vibrating string whose ends are immersed in a viscous liquid would satisfy (7) with $b > 0$ since energy is radiated to the liquid.

CONDITIONS AT INFINITY

In case the domain D is unbounded, the physics usually provides conditions at infinity. These can be tricky. An example is Schrödinger's equation, where the domain D is all of space, and we require that $\int |u|^2 \, d\mathbf{x} = 1$. The finiteness of this integral means, in effect, that u "vanishes at infinity."

A second example is afforded by the scattering of acoustic or electromagnetic waves. If we want to study sound or light waves that are radiating outward (to infinity), the appropriate condition at infinity is "Sommerfeld's outgoing radiation condition"

$$\lim_{r\to\infty} r\left(\frac{\partial u}{\partial r} - \frac{\partial u}{\partial t}\right) = 0, \tag{8}$$

where $r = |\mathbf{x}|$ is the spherical coordinate. (In a given mathematical context this limit would be made more precise.) (See Section 13.3.)

JUMP CONDITIONS

These occur when the domain D has two parts, $D = D_1 \cup D_2$ (see Figure 2), with different physical properties. An example is heat conduction, where D_1 and D_2 consist of two different materials (see Exercise 6).

EXERCISES

1. By trial and error, find a solution of the diffusion equation $u_t = u_{xx}$ with the initial condition $u(x, 0) = x^2$.
2. (a) Show that the temperature of a metal rod, insulated at the end $x = 0$, satisfies the boundary condition $\partial u/\partial x = 0$. (Use Fourier's law.)
   (b) Do the same for the diffusion of gas along a tube that is closed off at the end $x = 0$. (Use Fick's law.)

(c) Show that the three-dimensional version of (a) (insulated solid) or (b) (impermeable container) leads to the boundary condition $\partial u/\partial n = 0$.

3. A homogeneous body occupying the solid region D is completely insulated. Its initial temperature is $f(\mathbf{x})$. Find the steady-state temperature that it reaches after a long time. (*Hint:* No heat is gained or lost.)
4. A rod occupying the interval $0 \le x \le l$ is subject to the heat source $f(x) = 0$ for $0 < x < \frac{l}{2}$, and $f(x) = H$ for $\frac{l}{2} < x < l$ where $H > 0$. The rod has physical constants $c = \rho = \kappa = 1$, and its ends are kept at zero temperature.
   (a) Find the steady-state temperature of the rod.
   (b) Which point is the hottest, and what is the temperature there?
5. In Exercise 1.3.4, find the boundary condition if the particles lie above an impermeable horizontal plane $z = a$.
6. Two homogeneous rods have the same cross section, specific heat c, and density ρ but different heat conductivities κ_1 and κ_2 and lengths L_1 and L_2. Let $k_j = \kappa_j/c\rho$ be their diffusion constants. They are welded together so that the temperature u and the heat flux κu_x at the weld are continuous. The left-hand rod has its left end maintained at temperature zero. The right-hand rod has its right end maintained at temperature T degrees.
   (a) Find the *equilibrium* temperature distribution in the composite rod.
   (b) Sketch it as a function of x in case $k_1 = 2,\ k_2 = 1,\ L_1 = 3,\ L_2 = 2$, and $T = 10$. (This exercise requires a lot of elementary algebra, but it's worth it.)
7. In linearized gas dynamics (sound), verify the following.
   (a) If curl $\mathbf{v} = \mathbf{0}$ at $t = 0$, then curl $\mathbf{v} = \mathbf{0}$ at all later times.
   (b) Each component of $\mathbf{v}$ and ρ satifies the wave equation.

1.5 WELL-POSED PROBLEMS

Well-posed problems consist of a PDE in a domain together with a set of initial and/or boundary conditions (or other auxiliary conditions) that enjoy the following fundamental properties:

(i) *Existence:* There exists at least one solution $u(x, t)$ satisfying all these conditions.
(ii) *Uniqueness:* There is at most one solution.
(iii) *Stability:* The unique solution $u(x, t)$ depends in a stable manner on the data of the problem. This means that if the data are changed a little, the corresponding solution changes only a little.

For a physical problem modeled by a PDE, the scientist normally tries to formulate physically realistic auxiliary conditions which all together make a well-posed problem. The mathematician tries to prove that a given problem is or is not well-posed. If too few auxiliary conditions are imposed, then there may be more than one solution (*nonuniqueness*) and the problem is called underdetermined. If, on the other hand, there are too many auxiliary conditions, there may be no solution at all (*nonexistence*) and the problem is called overdetermined.

The stability property (iii) is normally required in models of physical problems. This is because you could never measure the data with mathematical precision but only up to some number of decimal places. You cannot distinguish a set of data from a tiny perturbation of it. The solution ought not be significantly affected by such tiny perturbations, so it should change very little.

Let us take an example. We know that a vibrating string with an external force, whose ends are moved in a specified way, satisfies the problem

$$\begin{aligned} &Tu_{tt} - \rho u_{xx} = f(x,t) \\ &u(x,0) = \phi(x) \qquad u_t(x,0) = \psi(x) \\ &u(0,t) = g(t) \qquad u(L,t) = h(t) \end{aligned} \tag{1}$$

for $0 < x < L$. The *data* for this problem consist of the five functions $f(x,t)$, $\phi(x)$, $\psi(x)$, $g(t)$, and $h(t)$. Existence and uniqueness would mean that there is exactly one solution $u(x, t)$ for arbitrary (differentiable) functions f, ϕ, ψ, g, h. Stability would mean that if any of these five functions are slightly perturbed, then u is also changed only slightly. To make this precise requires a definition of the "*nearness*" of functions. Mathematically, this requires the concept of a "distance", "metric", "norm", or "topology" in function space and will be discussed in the context of specific examples (see Sections 2.3, 3.4, or 5.5). Problem (1) is indeed well-posed if we make the appropriate choice of "nearness."

As a second example, consider the diffusion equation. Given an initial condition $u(x, 0) = f(x)$, we expect a unique solution, in fact, well-posedness, for $t > 0$. But consider the backwards problem! Given $f(x)$, find $u(x, t)$ for $t < 0$. What past behavior could have led up to the concentration $f(x)$ at time 0? Any chemist knows that diffusion is a smoothing process since the concentration of a substance tends to flatten out. Going backward ("*antidiffusion*"), the situation becomes more and more chaotic. Therefore, you would *not* expect well-posedness of the backward-in-time problem for the diffusion equation.

As a third example, consider solving a matrix equation instead of a PDE: namely, $Au = b$, where A is an $m \times n$ matrix and b is a given m-vector. The "data" of this problem comprise the vector b. If $m > n$, there are more rows than columns and the system is overdetermined. This means that no solution can exist for certain vectors b; that is, you don't necessarily have existence. If, on the other hand, $n > m$, there are more columns than rows and the system is underdetermined. This means that there are lots of solutions for certain vectors b; that is, you can't have uniqueness.

Now suppose that $m = n$ but A is a singular matrix; that is, $\det A = 0$ or A has no inverse. Then the problem is still ill-posed (neither existence nor uniqueness). It is also unstable. To illustrate the instability further, consider a nonsingular matrix A with one very small eigenvalue. The solution is unique but if b is slightly perturbed, then the error will be greatly magnified in the solution u. Such a matrix, in the context of scientific computation, is called ill-conditioned. The ill-conditioning comes from the instability of the matrix equation with a singular matrix.

As a fourth example, consider Laplace's equation $u_{xx} + u_{yy} = 0$ in the region $D = \{-\infty < x < \infty, 0 < y < \infty\}$. It is *not* a well-posed problem to specify both u and u_y on the boundary of D, for the following reason. It has the solutions

$$u_n(x, y) = \frac{1}{n} e^{-\sqrt{n}} \sin nx \sinh ny. \tag{2}$$

Notice that they have boundary data $u_n(x, 0) = 0$ and $\partial u_n/\partial y(x, 0) = e^{-\sqrt{n}} \sin nx$, which tends to zero as $n \to \infty$. But for $y \neq 0$ the solutions $u_n(x, y)$ do not tend to zero as $n \to \infty$. Thus the stability condition (iii) is violated.

EXERCISES

1. Consider the problem

$$\frac{d^2u}{dx^2} + u = 0$$
$$u(0) = 0 \quad \text{and} \quad u(L) = 0,$$

consisting of an ODE and a pair of boundary conditions. Clearly, the function $u(x) \equiv 0$ is a solution. Is this solution *unique, or not*? Does the answer depend on L?

2. Consider the problem

$$u''(x) + u'(x) = f(x)$$
$$u'(0) = u(0) = \tfrac{1}{2}[u'(l) + u(l)],$$

with $f(x)$ a given function.
   (a) Is the solution *unique*? Explain.
   (b) Does a solution necessarily *exist*, or is there a condition that $f(x)$ must satisfy for existence? Explain.

3. Solve the boundary problem $u'' = 0$ for $0 < x < 1$ with $u'(0) + ku(0) = 0$ and $u'(1) \pm ku(1) = 0$. Do the $+$ and $-$ cases separately. What is special about the case $k = 2$?

4. Consider the Neumann problem

$$\Delta u = f(x, y, z) \quad \text{in } D$$
$$\frac{\partial u}{\partial n} = 0 \quad \text{on bdy } D.$$

(a) What can we surely add to any solution to get another solution? So we don't have uniqueness.
(b) Use the divergence theorem and the PDE to show that

$$\iiint_D f(x, y, z)\, dx\, dy\, dz = 0$$

is a necessary condition for the Neumann problem to have a solution.
(c) Can you give a physical interpretation of part (a) and/or (b) for either heat flow or diffusion?

5. Consider the equation

$$u_x + yu_y = 0$$

with the boundary condition $u(x, 0) = \phi(x)$.
(a) For $\phi(x) \equiv x$, show that no solution exists.
(b) For $\phi(x) \equiv 1$, show that there are many solutions.

6. Solve the equation $u_x + 2xy^2u_y = 0$.

1.6 TYPES OF SECOND-ORDER EQUATIONS

In this section we show how the Laplace, wave, and diffusion equations are in some sense typical among all second-order PDEs. However, these three equations are quite different from each other. It is natural that the Laplace equation $u_{xx} + u_{yy} = 0$ and the wave equation $u_{xx} - u_{yy} = 0$ should have very different properties. After all, the *algebraic* equation $x^2 + y^2 = 1$ represents a circle, whereas the equation $x^2 - y^2 = 1$ represents a hyperbola. The parabola is somehow in between.

In general, let's consider the PDE

$$\boxed{a_{11}u_{xx} + 2a_{12}u_{xy} + a_{22}u_{yy} + a_1u_x + a_2u_y + a_0u = 0.} \qquad (1)$$

This is a linear equation of order two in two variables with six real constant coefficients. (The factor 2 is introduced for convenience.)

Theorem 1. By a linear transformation of the independent variables, the equation can be reduced to one of three forms, as follows.

(i) *Elliptic case:* If $a_{12}^2 < a_{11}a_{22}$, it is reducible to

$$u_{xx} + u_{yy} + \cdots = 0$$

(where $\cdots$ denotes terms of order 1 or 0).

(ii) *Hyperbolic case:* If $a_{12}^2 > a_{11}a_{22}$, it is reducible to

$$u_{xx} - u_{yy} + \cdots = 0.$$

(iii) *Parabolic case:* If $a_{12}^2 = a_{11}a_{22}$, it is reducible to

$$u_{xx} + \cdots = 0$$

(unless $a_{11} = a_{12} = a_{22} = 0$).

The proof is easy and is just like the analysis of conic sections in analytic geometry as either ellipses, hyperbolas, or parabolas. For simplicity, let's suppose that $a_{11} = 1$ and $a_1 = a_2 = a_0 = 0$. By completing the square, we can then write (1) as

$$(\partial_x + a_{12}\partial_y)^2 u + \left(a_{22} - a_{12}^2\right)\partial_y^2 u = 0 \tag{2}$$

(where we use the operator notation $\partial_x = \partial/\partial x$, $\partial_y^2 = \partial^2/\partial y^2$, etc.). In the elliptic case, $a_{12}^2 < a_{22}$. Let $b = (a_{22} - a_{12}^2)^{1/2} > 0$. Introduce the new variables ξ and η by

$$x = \xi, \quad y = a_{12}\xi + b\eta. \tag{3}$$

Then $\partial_\xi = 1 \cdot \partial_x + a_{12}\partial_y$, $\partial_\eta = 0 \cdot \partial_x + b\partial_y$, so that the equation becomes

$$\partial_\xi^2 u + \partial_\eta^2 u = 0, \tag{4}$$

which is Laplace's. The procedure is similar in the other cases. □

Example 1.

Classify each of the equations
(a) $u_{xx} - 5u_{xy} = 0$.
(b) $4u_{xx} - 12u_{xy} + 9u_{yy} + u_y = 0$.
(c) $4u_{xx} + 6u_{xy} + 9u_{yy} = 0$.

Indeed, we check the sign of the "discriminant" $\mathscr{D} = a_{12}^2 - a_{11}a_{22}$. For (a) we have $\mathscr{D} = (-5/2)^2 - (1)(0) = 25/4 > 0$, so it is hyperbolic. For (b), we have $\mathscr{D} = (-6)^2 - (4)(9) = 36 - 36 = 0$, so it is parabolic. For (c), we have $\mathscr{D} = 3^2 - (4)(9) = 9 - 36 < 0$, so it is elliptic. □

The same analysis can be done in any number of variables, using a bit of linear algebra. Suppose that there are n variables, denoted $x_1, x_2 \ldots, x_n$, and the equation is

$$\sum_{i,j=1}^{n} a_{ij}u_{x_i x_j} + \sum_{i=1}^{n} a_i u_{x_i} + a_0 u = 0, \tag{5}$$

with real constants a_{ij}, a_i, and a_0. Since the mixed derivatives are equal, we may as well assume that $a_{ij} = a_{ji}$. Let $\mathbf{x} = (x_1, \ldots, x_n)$. Consider any linear change of independent variables:

$$(\xi_1, \ldots, \xi_n) = \boldsymbol{\xi} = B\mathbf{x},$$

where B is an $n \times n$ matrix. That is,

$$\xi_k = \sum_m b_{km}x_m. \tag{6}$$

Convert to the new variables using the chain rule:

$$\frac{\partial}{\partial x_i} = \sum_k \frac{\partial \xi_k}{\partial x_i} \frac{\partial}{\partial \boldsymbol{\xi}_k}$$

and

$$u_{x_i x_j} = \left(\sum_k b_{ki} \frac{\partial}{\partial \xi_k}\right)\left(\sum_l b_{lj} \frac{\partial}{\partial \xi_l}\right) u.$$

Therefore the PDE is converted to

$$\sum_{i,j} a_{ij} u_{x_i x_j} = \sum_{k,l}\left(\sum_{i,j} b_{ki} a_{ij} b_{lj}\right) u_{\xi_k \xi_l}. \tag{7}$$

(Watch out that on the left side u is considered as a function of $\mathbf{x}$, whereas on the right side it is considered as a function of $\boldsymbol{\xi}$.) So you get a second-order equation in the new variables $\boldsymbol{\xi}$, but with the *new coefficient matrix* given within the parentheses. That is, the new matrix is

$$BA^tB,$$

where $A = (a_{ij})$ is the original coefficient matrix, the matrix $B = (b_{ij})$ defines the transformation, and ${}^tB = (b_{ji})$ is its transpose.

Now a theorem of linear algebra says that for any symmetric real matrix A, there is a rotation B (an orthogonal matrix with determinant 1) such that BA^tB is the diagonal matrix

$$BA^tB = D = \begin{pmatrix} d_1 & & & & & \\ & d_2 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & d_n \end{pmatrix}. \tag{8}$$

The real numbers $d_1, \ldots, d_n$ are the eigenvalues of A. Finally, a change of scale would convert D into a diagonal matrix with each of the d's equal to $+1$, -1, or 0. (This is what we did, in effect, early in this section for the case $n = 2$.)

Thus any PDE of the form (5) can be converted by means of a linear change of variables into a PDE with a diagonal coefficient matrix.

Definition. The PDE (5) is called *elliptic* if all the eigenvalues $d_1, \ldots, d_n$ are positive or all are negative. [This is equivalent to saying that the original coefficient matrix A (or $-A$) is positive definite.] The PDE is called *hyperbolic* if none of the $d_1, \ldots, d_n$ vanish and one of them has the opposite sign from the $(n - 1)$ others. If none vanish, but at least two of them are positive and at least two are negative, it is called *ultrahyperbolic*. If exactly one of the eigenvalues is zero and all the others have the same sign, the PDE is called *parabolic*.

Ultrahyperbolic equations occur quite rarely in physics and mathematics, so we shall not discuss them further. Just as each of the three conic sections has quite distinct properties (boundedness, shape, asymptotes), so do each of the three main types of PDEs. □

More generally, if the coefficients are variable, that is, the a_{ij} are functions of $\mathbf{x}$, the equation may be elliptic in one region and hyperbolic in another.

Example 2.

Find the regions in the xy plane where the equation

$$yu_{xx} - 2u_{xy} + xu_{yy} = 0$$

is elliptic, hyperbolic, or parabolic. Indeed, $\mathscr{D} = (-1)^2 - (y)(x) = 1 - xy$. So the equation is parabolic on the hyperbola ($xy = 1$), elliptic in the two convex regions ($xy > 1$), and hyperbolic in the connected region ($xy < 1$). □

If the equation is nonlinear, the regions of ellipticity (and so on) may depend on which solution we are considering. Sometimes nonlinear transformations, instead of linear transformations such as B above, are important. But this is a complicated subject that is poorly understood.

EXERCISES

1. What is the type of each of the following equations?
   (a) $u_{xx} - u_{xy} + 2u_y + u_{yy} - 3u_{yx} + 4u = 0$.
   (b) $9u_{xx} + 6u_{xy} + u_{yy} + u_x = 0$.
2. Find the regions in the xy plane where the equation

$$(1 + x)u_{xx} + 2xyu_{xy} - y^2u_{yy} = 0$$

   is elliptic, hyperbolic, or parabolic. Sketch them.
3. Among all the equations of the form (1), show that the only ones that are unchanged under all rotations (*rotationally invariant*) have the form $a(u_{xx} + u_{yy}) + bu = 0$.
4. What is the *type* of the equation

$$u_{xx} - 4u_{xy} + 4u_{yy} = 0?$$

   Show by direct substitution that $u(x, y) = f(y + 2x) + xg(y + 2x)$ is a solution for arbitrary functions f and g.
5. Reduce the elliptic equation

$$u_{xx} + 3u_{yy} - 2u_x + 24u_y + 5u = 0$$

to the form $v_{xx} + v_{yy} + cv = 0$ by a change of dependent variable $u = ve^{\alpha x + \beta y}$ and then a change of scale $y' = \gamma y$.

6. Consider the equation $3u_y + u_{xy} = 0$.
   (a) What is its type?
   (b) Find the general solution. (*Hint:* Substitute $v = u_y$.)
   (c) With the auxiliary conditions $u(x, 0) = e^{-3x}$ and $u_y(x, 0) = 0$, does a solution exist? Is it unique?