Introduction and some preliminaries

1 Partial differential equations

A *partial differential equation* (PDE) is a relationship among partial derivatives of a function (or functions) of more than one variable. In contrast, ordinary differential equations have only one independent variable. Problems of this form are often supplemented with side conditions on the domain boundaries.

The study of PDEs is quite old (at least 250 years) and is still evolving. Its development is intertwined with the very fabric of modern scientific discovery. This is particularly true of physics, since most of the elementary (and not so elementary) PDEs were developed from descriptions of fluid and solid mechanics, electromagnetism, and quantum mechanics. More recently, the subject has filtered into other scientific endeavors, such as ecology (e.g. the Fisher equation), chemistry (e.g. reaction diffusion equations), materials science (e.g. the Cahn-Hilliard equation). Contemporary uses of PDEs have grown ever more imaginative, spanning topics from image processing to mathematical finance. Some famous examples are shown in table 1. The first three comprise the most basic linear equations; we shall study them first.

Name	Equation	Applications
Wave equation	$u_{tt} = c^2 u_{xx}$	Vibrating string
Diffusion equation	$u_t = D u_{xx}$	Heat flow
Laplace's equation	$u_{xx} + u_{yy} = 0$	Electrostatics and many more
Burger's equation	$u_t + u u_x = D u_{xx}$	Fluid mechanics
Cahn-Hilliard equation	$u_t = (u^3 - u - u_{xx})_{xx}$	Phase separation
Eikonal equation	$\lvert\nabla u\rvert = f(x)$	Waves and control theory
Fisher's equation	$u_t = u_{xx} + u(1-u)$	Ecology
Korteweg-de Vries equation	$u_t + 6uu_x + u_{xxx} = 0$	Water waves
Schrödinger equation	$iu_t + u_{xx} = 0$	Quantum mechanics
Nonlinear Schrödinger equation	$iu_t + u_{xx} + \lvert u\rvert^2 u = 0$	Nonlinear optics
Swift-Hohenberg equation	$u_t = -(\partial_{xx} + 1)^2 u + N(u)$	Pattern formation

Table 1: Some of the famous PDEs

The simplest classification of PDE's is *order*, which is the highest derivative that appears. Often, there is a distinction between order in space and time. For example, the wave equation is second order in both x and t derivatives, whereas the diffusion equation is first order in the time variable t but second order in the spatial variable x.

1.1 Coordinates and domains

In many problems, we often associate independent variables to coordinates of a physical system. For example, a function $u(x, y, z, t)$ might be regarded as assigning a value to each point $(x, y, z) \in \mathbb{R}^3$ at time t. It is not always essential to impart a physical meaning to variables, but the identification of spatial variables is important for defining operations like gradient, divergence and Laplacian (see below). It is also nice to have a physical interpretation of an otherwise abstract functional relation.

The domain of the function(s) involved is an an essential aspect in the study of PDEs. Here the domain will typically be denoted Ω, and will be regarded as an open set (one that does not contain its boundary points). An equation is expected to be satisfied within the domain Ω, but nowhere else.

The coordinate system we use typically reflects the type of domain. For example a function $u : \mathbb{R}^2 \to \mathbb{R}$ might be written in Cartesian coordinates $u(x, y)$ for a square domain or polar variables $u(r, \theta)$ for a semi-circular domain. Of course, a PDE written in terms of $x-$ and $y-$ partial derivatives is going to be different in terms of r and $\theta-$ partial derivatives. There is a perfectly straightforward (although often laborious) way of changing coordinates in a PDE by the chain rule. For example, we can regard polar coordinates as functions of Cartesian ones as $r = r(x, y)$ and $\theta = \theta(x, y)$. Thus the partial derivative

$$u_x = u_r r_x + u_\theta \theta_x.$$

Of course, to effect a complete change of variables, one needs to write r_x and θ_x in terms of r and θ alone.

Sometimes there will be no need to refer to a particular coordinate system. In this case, the spatial coordinate is identified by a (boldface) vector $\mathbf{x}$. In fact, many calculations in PDE are far simpler to accomplish using this so-called "coordinate-free" notation.

1.2 Side conditions

A PDE alone is generally insufficient to describe a situation of interest. Often there are *side conditions* which are to be satisfied in addition to the equation itself. These are typically *boundary* or *initial* conditions, which as their names imply restrict the behavior of the solution on the physical boundary or at some specified value of time, respectively.

There are two common types of boundary conditions encountered in applications. The first is called *Dirichlet*, which simply fixes the value of the solution on the boundary, e.g.

$$u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega. \tag{1}$$

The notation $\partial\Omega$ refers to the boundary of the set Ω. A second frequently encountered type of boundary condition fixes the derivative normal to the boundary

$$\frac{\partial u}{\partial n} = \nabla u(\mathbf{x}) \cdot \hat{\mathbf{n}} = f(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega, \tag{2}$$

where by convention $\hat{\mathbf{n}}$ is a unit vector normal to $\delta\Omega$ and points outward (assuming the boundary is a smooth curve or surface). Condition (1) is often called the *Neumann* condition. For either case (1) or (2), if $f \equiv 0$, these boundary conditions are called *homogeneous*.

In many problems, the domain is an unbounded set such as $\mathbb{R}^n$ or $\{(x, y), y > 0\}$. In these cases, it is useful to think of the domain boundary as it were moved out to infinity. Rather than a boundary condition for "boundaries at infinity", we have *far-field conditions*, which specify some kind of limiting behavior of the solution. An example of this is

$$\lim_{|\mathbf{x}| \to \infty} u(\mathbf{x}) = 0,$$

which behaves as a sort of Dirichlet condition.

For problems where time (or something like it) is one of the independent variables, initial conditions are sometimes imposed. For example, one might specify the initial value of a solution

$u(x,t)$ as $u(x,0) = g(x)$ or even the initial time derivative, for example $u_t(x,0) = h(x)$. Although the notion of an initial condition is probably familiar from ordinary differential equations, here its value is not a number, but a function of the spatial variables.

It will be occasionally useful to ignore or change certain side conditions when finding solutions to PDEs. The expectation is that when fewer conditions are imposed on a problem, it will possess more solutions. This has the advantage of generating *families* of solutions, from which the relevant ones may be extracted.

1.3 Well-posedness

What makes for a "good" math problem? [1] Usually when we model the physical world, we expect our equations to reproduce what we observe, and good science demands that our observations are themselves reproducible. So if our model is complete (which is sometimes too strong of a demand), it should give us one and only one solution, which hopefully mimics the real world observation we were trying to model.

In the context of differential equations, *well-posed* problems have the following properties.

- Existence: there is at least one solution to equation and side conditions. If this were not the case, the model is perhaps too restrictive, and may not tell us anything!
- Uniqueness: there is only one solution. This property goes back to the desire for reproducibility: if we get many solutions, which one would we expect to observe? It often happens models do not have unique solutions, and it is generally because of missing ingredients such as side conditions.
- Stability: informally speaking, this means if the model problem is changed a little, then the solution only changes a little. Since the inputs (side conditions, coefficients, etc.) to models are generally only known approximately, it would be undesirable if the solution was highly sensitive to these things.

2 Multivariable derivatives, integrals and notation

The reader should be familiar with the basics - vector calculus and ordinary differential equations in particular. The following section merely mentions some of the highlights from calculus of several variables, and establishes the particular notation which will be employed throughout.

2.1 Derivatives of more than one variable

Functions with more than one independent variable have various kinds of derivatives: partial derivatives, directional derivatives, gradients, divergence, curl, etc. The most important one is the partial derivative

$$\frac{\partial}{\partial x}u(x,y) = \frac{d}{dx}\Big|_y u(x,y)$$

where the notation $|_y$ means "treat y as if it were fixed". We will typically use a more compact notation $\partial_x u$ or u_x for partial derivatives (the former is useful when talking about the derivative

[1] The cynic would say, anything that gets you funding!

as an operator). Multiple partial derivatives are notated similarly, e.g.

$$\frac{\partial^2 u}{\partial x \partial y} = u_{xy} = \partial_{xy} u.$$

Recall that if u has continuous derivatives, the order of differentiation does not matter, so $u_{xy} = u_{yx}$. A more general notion of partial derivatives comes from specifying a direction $\mathbf{d}$ in which the derivative is to taken. Provided $\mathbf{d}$ has unit length, then the directional derivative in this direction is $\nabla u \cdot \mathbf{d} = \partial u / \partial d$. For example the derivative in (2) is the derivative in the direction of the normal vector.

We occasionally need to make use of the chain rule, e.g.

$$\frac{\partial}{\partial r} u(x(r,s), y(r,s)) = u_x(x(r,s), y(r,s)) x_r + u_y(x(r,s), y(r,s)) y_r.$$

Recall that the idea behind this is simple: just add all possible combinations of one-dimensional chain rules which link the function and the independent variable.

Often conceptual trouble with partial derivatives comes in situations where variables are used ambiguously both to name arguments (inputs) of a function as well as label variables which may comprise the inputs. For example, suppose you have some function of two variables $w(x,t)$ and want to compute $\partial_t w(x - ct, t)$. Does this mean fix x and differentiate with respect to t, or fix the first input of w and differentiate with respect to the second variable? To clear up this confusion, it is best to temporarily rename the arguments of w, to (say) $w = w(y,s)$, where $y = x - ct$ and $s = t$. Then using the chain rule,

$$\partial_t w(x - ct, t) = w_y y_t + w_s s_t = -c w_y(x - ct, t) + w_s(x - ct, t)$$

which is equivalent in original variables to $-c w_x(x - ct, t) + w_t(x - ct, t)$.

We shall use other types of derivatives, notably gradient and divergence. These operations are generally only meant to apply only to the spatial variables, so by convention

$$\nabla u(x,y,t) = [u_x, u_y], \quad \nabla \cdot [g(x,y,t), h(x,y,t)] = g_x + h_y,$$

where vectors are denoted as ordered pairs $[x_1, x_2]$ In the context of PDEs, there is a special differential operator which arises frequently, the *Laplace* operator (or simply Laplacian), which is simply the divergence of the gradient. In three dimensions, it is notated [2] as

$$\nabla \cdot [\nabla u(x,y,z,t)] = u_{xx} + u_{yy} + u_{zz} \equiv \Delta u.$$

We will study this operator extensively.

2.2 Integration over areas, volumes, curves, surfaces

Multidimensional integration is often encountered in the modeling, analysis and solution of PDEs. We adopt a generic notation for integrals over a region R of a function whose domain contains R,

$$\int_R f(\mathbf{x}) d\mathbf{x}. \tag{3}$$

Implicit is the fact f is a function of a vector coordinate $\mathbf{x}$. The region R can be a variety of objects - areas, volumes, curves, surfaces. The notation $d\mathbf{x}$ simply means integrate with respect to the

[2] In physics, the Laplacian is unfortunately written ∇^2, even though it is NOT the gradient of a gradient.

variable(s) denoted by the vector $\mathbf{x}$. This is important since the integrand may be a function of other variables. For example, it might be that $f = f(\mathbf{x}_1, \mathbf{x}_2)$ is a function of two sets of vectors $\mathbf{x}_1$ and $\mathbf{x}_2$, but we only want to integrate with respect to one of them, $\int_R f(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1$. We treat $\mathbf{x}_2$ like a constant in this case, and only $\mathbf{x}_1$ ranges over values in the set R.

It is (almost) always better to notate multidimensional integrals like (3) without referring to a particular coordinate system. When it comes to actually computing these integrals, however, we sometimes resort to writing everything with respect to coordinates chosen to be convenient (Cartesian, polar, etc.) and parameterizing R in terms of these coordinates. This usually results in an iterated integral like

$$\int_a^b \int_c^d \int_e^f \cdots f(x, y, z, \ldots) dx\, dy\, dz \ldots$$

Such an integral is often mistakenly called a volume integral, but it is not - it is simply an iterated set of plain one dimensional integrals which happens to represent the multidimensional integral (3).

A special case of a integral over a curve or surface arises from using the divergence theorem

$$\int_R \nabla \cdot \mathbf{F}(\mathbf{x}) d\mathbf{x} = \int_{\partial R} \mathbf{F}(\mathbf{x}) \cdot \hat{\mathbf{n}}(\mathbf{x}) d\mathbf{x}.$$

The region R is a closed area (in $\mathbb{R}^2$) or volume (in $\mathbb{R}^3$), and ∂R denotes the boundary of R. The integral on the right is usually called a *flux* integral, since it involves the flow of the vector field $\mathbf{F}$ in the direction $\hat{\mathbf{n}}(\mathbf{x})$ normal to the boundary. Since the normal vector $\hat{\mathbf{n}}$ depends on the point on the boundary, it is often necessary to notate it as a function of $\mathbf{x}$ (especially if the integrand happens to depend on two sets of coordinates).