CHAPTER 1
Introduction

1.1 GENERAL NATURE OF PDE

It is no exaggeration to state that partial differential equations (PDE) have played a vital role in the development of science and technology, primarily since the beginning of the twentieth century. In the earlier stage, PDE were mainly used to describe physical phenomena, like vibrations of strings, heat conduction in solids, transport phenomena, to mention a few. Later, with the advantage of mathematical modelling, the scope of using PDE for the description of phenomena occurring in biology, economics and even sociology became prominent.

Since the days of Newton or even earlier, many have attempted to describe physical processes using mathematics.[1] Such a mathematical description often leads to linear differential, integral and even integro-differential equations. Thus, a large number of PDE naturally come from mathematical physics. The initial developments in PDE, though, were mainly geared towards obtaining solutions to a particular physical or engineering problem, it was soon realized that many of the problems will have common features and similarities. This naturally led to the grouping of PDE that can be tackled in a single framework. This automatically leads to the abstraction of the subject and the theoretical analysis that follows, hence, becomes more important. This is one of the features we try to follow in the present book. Indeed, unlike ordinary differential equations (ODE), all PDE including the linear ones cannot be treated in a single theoretical framework, leading to the necessity of a classification. In fact, due to the diverse nature of physical phenomena, we remark that we cannot classify all the PDE. Nevertheless, a fairly good classification is available for the second-order equations and interestingly a large number of physical and other problems lead to second-order equations. Also, for the three important classes of equations, namely elliptic, hyperbolic and parabolic, general theories have been developed.

As mentioned above, a wide class of physical problems is described by second-order linear differential equations of the form

$$\sum_{i,j=1}^{n} a_{ij}(x)u_{x_i x_j} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u = f(x). \tag{1.1}$$

[1] It is a historical fact that the *calculus* was born during such a process.

Here the variable x varies in an open set in the physical space $\mathbb{R}^n$, $n = 1, 2, 3$ and the coefficients a_{ij}, b_i and c are known from the physical process; u is the unknown function and f denotes an external quantity, if any, influencing the physical process.

We only mention a few real-world situations where PDE occur. For more examples and their detailed discussion, the reader is referred to Barták et al. (1991), Markowitz (2005), Murray (2003), Rhee et al. (1986), and Vladimirov (1984).

Many problems in mechanics like vibrations of strings, rods, membranes and three-dimensional objects and also the mathematical description of electromagnetic waves lead to the equation of vibrations, which is the wave equation in one more space dimension. If the mean free path of the particles is much larger than their dimensions, then the propagation of a particle may be more accurately described by an equation, in comparison with the diffusion equation, called the *transport* or *kinetic equation*. This is also called the *Boltzmann equation*. This is an integro-differential equation.

The Heisenberg principle states that the position of a particle and its momentum cannot be *simultaneously* described, according to the laws of quantum mechanics. Thus, for example, the position of a quantum particle can be confirmed only with certain probability. The Schrödinger's equation is an attempt to describe the dynamics of a quantum particle of a given mass moving in an external force field with a given potential. Reaction–Diffusion equations describe the interaction of two or more chemical concentrations of distinct diffusivity coefficients, in a chemical or biological process. These equations are also used in the modelling of pattern formation and form an important part of *Mathematical Biology* and constitute a system of non-linear diffusion equations.

The equation of heat diffusion in a medium and the diffusion of a chemical species are described by the heat or diffusion equation. Euler's equations of gas dynamics describe the dynamics of an ideal fluid, that is, a fluid with no or negligible viscosity. These equations form a system of first-order hyperbolic equations. In a particular situation where liquid is incompressible and has a potential, these equations reduce to the Poisson's equation for the potential function. The system of Maxwell's equations describe the dynamics of a charged particle in a medium with varying electromagnetic field, invoking *Ampere's law* and *Faraday's law*. In some particular cases, each component of the electric and magnetic fields satisfies the *telegraph equation*.

1.2 TWO EXAMPLES

The following two situations perhaps describe a general nature in the analysis of solutions to PDE. These are quite simple to state and involve second-order equations in two variables. The equations are the Laplace equation, the heat or diffusion equation and the wave equation: $L_i u = 0,\ i = 1, 2, 3$, where

$$L_1 = \partial_t^2 + \partial_x^2,\ L_2 = \partial_t - \partial_x^2 \text{ and } L_3 = \partial_t^2 - \partial_x^2.$$

The first situation involves the determination of solutions of $L_i u = 0$, $i = 1, 2, 3$ with prescribed data on the boundary of a rectangle $ABCD$ with the side AB situated on the x-axis in the $x-t$ plane. Without much concern whether to prescribe u or its first derivatives on the sides of $ABCD$, let us dwell on the number of conditions required for each of the operators L_i, in order to determine a solution of $L_i u = 0$. It turns out that L_1 requires *four* conditions one each on the four sides of the rectangle $ABCD$; L_2 requires *three* conditions one each on the sides AB, BC and AD of the rectangle $ABCD$; L_3 requires *four* conditions – two on AB and one each on BC and AD of the rectangle $ABCD$.

Note that all the three operators are linear and of second order. Yet, the number of data to be prescribed and the part of the boundary where to be placed become important in order to determine a solution. Apparently, there is no simple explanation for this anomaly. Perhaps the reader will find an answer after studying the relevant chapters in the book. This is quite different from the analysis of an initial value problem (IVP) of a system of linear ODE; here the problem can be studied for a system of *any order* in a single framework. However, in the case of PDE, as the above examples exhibit, it is not possible to do an analysis even for second-order linear equations in two dimensions, in a single framework. This leads to the notion of a classification of PDE, and a particular condition on the data like initial values or boundary values depends on the type of PDE under consideration.

The second situation also concerns the operators L_i, but now with regard to *weak solutions* of them. A continuous or a locally integrable function u defined in an open set Ω in $\mathbb{R}^2$ is said to be a *weak solution* of $L_i u = 0$ for $i = 1, 2, 3$, if $\iint u(x,t) L_i \varphi(x,t)\, dxdt = 0$ for all $\varphi \in C_c^\infty(\Omega)$.

It is shown in Chapter 9 that any continuous or locally integrable function u of the form $v(x \pm t)$ is a weak solution for L_3 and thus it can admit discontinuous (weak) solutions. For the operators L_1 and L_2, it turns out that any weak solution is in fact a C^∞ function, may be after making corrections in a *set of measure zero*. The apparently strange behavior of the operators L_1 and L_2 cannot be explained in simple terms and the reader will not find a complete answer in this book! The operators L_1 and L_3 are quite different, but the operator L_2 may share some properties with L_1 (regularity) and some other properties with L_3 (energy estimates).

The above two situations describe, we hope, the complexities that are involved in the analysis of PDE. There is indeed constant evolution of the subject as and when some peculiar phenomenon is observed through an example or otherwise. In this connection, it is an interesting fact that a somewhat *true* picture of linear operators started emerging only after the work of Peetre (1960), even though there were already quite many advancements in the modern theory of PDE which had emerged through the works of Leray, Petrowski, Schwartz and others. With appropriate domain and range of the operator, what Peetre showed was that the linear operators are precisely the *local operators*. This means that $\mathrm{supp} Pu \subset \mathrm{supp} u$, where P is linear and u is in its domain. This then led to the discovery of *pseudo-differential operators* and *Fourier integral operators*. Roughly speaking, the *inverse* of an elliptic operator is a pseudo-differential operator and the *inverse* of a wave operator is a Fourier integral operator (see Nirenberg, 1976).

1.3 DESCRIPTION OF THE CONTENTS

This then sets the stage for the present book, with a modest list of contents.

- The first chapter briefly discusses certain general notions of PDE, their occurrence in physical and other sciences and engineering. It also describes the contents of the book, chapter-wise.
- The theory[2] of modern PDE is quite vast and demands a great amount of prerequisites such as Lebesgue integration theory, functional analysis, distributions and Sobolev spaces. Since we are discussing mostly classical theory in the present book, the prerequisites are minimal – a good understanding of multivariable calculus should suffice for studying this book. Exceptions do occur in Chapters 4 and 5, where the reader is expected to have a good knowledge of the modern theory of integration, especially in the proofs of uniqueness of solutions. In Chapter 2, we collect a good number of results from multivariable calculus, ODE and related topics that are used in the book. To make the book as self-contained as possible, we have also provided the proofs when they are not too lengthy.
- Chapter 3 is about the first-order equations. Here we study the general Cauchy problem (IVP) for such equations. The (local) theory is fairly complete as the problem is reduced to an IVP for a system of ODE. The geometry, however, does get complicated as we move from linear to quasilinear to general first-order equations. Because of their importance in applications, we mention two classes of first-order equations, namely the conservations laws and the Hamilton–Jacobi equations. These two classes are studied in detail in further chapters.
- In Chapters 4 and 5, we consider certain important class of first-order equations – Hamilton–Jacobi Equation (HJE) and Conservation Laws (CL) – which have been topics of great interest among researchers owing to their importance in many applications. Though these equations have been mentioned in Chapter 3, the emphasis here is on a new concept of a solution of these equations. A beginner perhaps encounters for the first time the concept of a *weak* solution to a PDE, which is in general a non-differentiable function! Furthermore, to obtain uniqueness of a solution, additional condition(s) need to be imposed. Since the theory of modern PDE largely deals with *weak* solutions, we thought it is a good idea to introduce this concept of solution to a beginner in the context of HJE and CL. However, these chapters may be skipped for the first reading as the uniqueness results require a good knowledge of modern theory of integration.
- In the context of ODE, the theory dealing with the Cauchy problem of a single equation or that of a system of first-order equations is essentially the same. In particular, the

[2] This is not to suggest there is a *single* theory of PDE, like theory of ODE or theory of functions of real or complex variable. In fact, we see in the literature different *theories* of PDE owing to the sheer vastness of the subject.

analysis is the same for both the first-order equations and higher-order equations, in the study of ODE. In contrast, such is not the situation about PDE. This makes the subject of PDE more complicated and also interesting. In Chapter 6, we explain how the data in a Cauchy problem for a second-order equation cannot in general be arbitrary. This naturally leads to the concept of classification of second- and higher-order equations. The main discussion in this chapter is about second-order equations and their solutions.

It should, however, be noted that some important developments in science in nineteenth and twentieth centuries, especially quantum mechanics and fluid dynamics, resulted in new types of PDE – the Schrödinger equation, Navier–Stokes equations and Kortweg-de Vries (K-dV) equation, for example. These equations and many more equations do not fall in the ambit of the above-mentioned classification. Thus, there were attempts to make the subject of PDE a unified subject without mentioning the class to which a PDE belongs. However, such attempts have not been that successful. This is one of the reasons we see a great number of books written on a particular equation or on a particular class of equations.

- Undoubtedly, the three major equations of mathematical physics – the Laplace equation (Poisson equation), the heat or diffusion equation, and the wave equation – have had great impact on the development of much of the modern theory of PDE. These equations are the topics of discussion in Chapters 7 through 10, respectively.
- The Laplace operator is a prototype of uniformly elliptic operators. Some important properties – mean value property, maximum (minimum) principle, Harnack's inequalities – enjoyed by a solution of the Laplace's equation are discussed at length in Chapter 7. We have also indicated that the maximum (minimum) principle is also enjoyed by a solution of a general uniformly elliptic equation. The existence and uniqueness of the solutions are also discussed via Perron's method and Newtonian potential.
- In Chapter 8, the heat equation and its solutions are studied in great detail. This equation is a prototype of parabolic equations. In a way this equation *sits* between the Laplace's equation and the wave equation. Therefore, its solution enjoys certain properties from both sides. For example, maximum (minimum) principle from Laplace's equation and energy estimate from the wave equation. Its solution also enjoys a mean value property and backward uniqueness property.
- The study of Laplace's equation and the heat equation largely does not depend on the dimension. However, the analysis of the wave equation does depend on the dimension and this is the reason to consider the study of the wave equation in one dimension and higher dimensions separately. These are dealt with in Chapters 9 and 10, respectively. The wave equation is a prototype of hyperbolic equations.
- The Cauchy–Kovalevsky theorem is, historically, an important result in the subject field of PDE. It is one of the first results proving the existence and uniqueness of solution to a Cauchy problem for a general equation, though in a restricted class of

equations with analytic coefficients. Nevertheless, the contents of its proof are full of a priori estimates, a hallmark of the modern theory of PDE. In Chapter 11, we present the details of this theorem and a generalization. We also briefly discuss the Holmgren's uniqueness result.

- We also briefly mention some aspects of the modern theory without going into details in Chapter 12. An existence result of L^2 weak solution is discussed here, to give a general flavor of a modern theory.

CHAPTER 2
Preliminaries

2.1 MULTIVARIABLE CALCULUS

2.1.1 Introduction

We plan to briefly introduce the calculus on $\mathbb{R}^n$, namely the concept of total derivative of multivalued function, $f = (f_1, \cdots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$. We are indeed familiar with the notion of partial derivatives $\partial_i f_j = \frac{\partial f_j}{\partial x_i}$, $1 \le i \le n, 1 \le j \le m$. In the sequel, we will introduce the important concept of *total derivative* and discuss its connection to the partial derivatives. We remark that the total derivative (known also as *Frechét derivative*) can be extended to infinite dimensional normed linear spaces, which is used in the analysis of more complicated problems especially arising from optimal control problems, calculus of variations, partial differential equations, and so on.

Motivation: One of the fundamental problems in mathematics (and hence in applications as well) is the following: Let $f : \mathbb{R}^n \to \mathbb{R}^n$. Given $y \in \mathbb{R}^n$, solve the system of equations

$$f(x) = y \tag{2.1}$$

and represent the solution as $x = g(y)$ and if possible find good properties of g, namely its smoothness. More generally, if $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$, $x \in \mathbb{R}^n, y \in \mathbb{R}^m$, solve the implicit system of equations

$$f(x, y) = 0 \tag{2.2}$$

and represent the solution as $x = g(y)$. Consider the one-dimensional case, where $f : \mathbb{R} \to \mathbb{R}$ which is C^1. Suppose that $f'(a) \neq 0$ for some a. Then, by the continuity of f', we see that $f'(x) \neq 0$ in a neighborhood interval I of a. Hence f' preserves the sign in I, f is monotonic in I and $f(I)$ is an interval. Thus, if $f(a) = b$, then the above argument shows that $f(x) = y$ is solvable for all y in $f(I)$, a neighborhood of b. This is the local solvability that is obtained by the non-vanishing property of the derivative of f at a. This immediately shows the importance of understanding the derivatives in the solvability of algebraic equations. We remark that the mere existence of all partial derivatives does not guarantee the local solvability. We need the stronger concept of total derivative.

Linear Systems: Let us look at the well-known linear system

$$Ax = y, \tag{2.3}$$

where $A = [a_{ij}]$ is a given $n \times n$ matrix. That is $f(x) = Ax$. The system (2.3) can be rewritten as

$$\sum_{j=1}^{n} a_{ij}x_j = y_i, \quad 1 \leq i \leq n. \tag{2.4}$$

The system (2.3) or (2.4) is uniquely solvable for x in terms of y if and only if $\det A \neq 0$ (global solvability). In this case

$$x = A^{-1}y$$

and A^{-1} is also an $n \times n$ matrix. We would like to address the solvability of (2.1) and (2.2) giving appropriate conditions similar to non-vanishing determinant as in the case of a linear system.

Example 2.1. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Clearly $f(0) = 0$. For $y > 0$, the equation $x^2 = y$ has two solutions $x_1 = +\sqrt{y}$ and $x_2 = -\sqrt{y}$ (non-uniqueness) and $y < 0$, the equation has no solution. Thus, we sense a difficulty around $x = 0$. Note that $\frac{\partial f}{\partial x}\Big|_{x=0} = 2x|_{x=0} = 0$. This shows that we cannot decide the sign of $\frac{\partial f}{\partial x}$ around 0. If we take any $a \neq 0$, and $b = a^2 = f(a)$, then, for any $y \in (b - \varepsilon, b + \varepsilon)$, ε small, there exists unique $x \in (a - \delta, a + \delta)$ for some δ such that $f(x) = y$. That is, the equation is solvable in a neighborhood of the point $b = f(a)$. Here, observe that $\frac{\partial f}{\partial x}\Big|_{x=a} = 2x|_{x=a} = 2a \neq 0$ and thus the sign of $\frac{\partial f}{\partial x}(a)$ is known. □

Example 2.2. Consider the function $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ defined by $f(x, y) = x^2 + y^2 - 1$. Indeed, the solutions (x, y) of the equation $f(x, y) = 0$ are points on the unit circle. Consider the solvability of x in terms of y near the solution $(0, 1)$ of $x^2 + y^2 - 1 = 0$, that is $x^2 = 1 - y^2$.

For y near 1, $y < 1$, we have two solutions $x_1 = +\sqrt{1 - y^2}$, $x_2 = -\sqrt{1 - y^2}$. Similarly the case near the point $(0, -1)$. Again observe that $\frac{\partial f}{\partial x}|_{(0,\pm1)} = 2x|_{(0,\pm1)} = 0$.

On the other hand, consider the point $(+1, 0)$. For y near 0, there exists unique solution $x = +\sqrt{1 - y^2}$; and for the point $(-1, 0)$ and y near 0, there exists unique solution $x = -\sqrt{1 - y^2}$. In fact, for any (a, b) with $a^2 + b^2 - 1 = 0$ and $a \neq 0$, we get $\frac{\partial f}{\partial x}\Big|_{(a,b)} \neq 0$ and the system is uniquely solvable for x in terms of y in a neighborhood of b. The situation is reversed if we look at the possibility of solving y in terms of x. □

Thus, we see the impact of non-vanishing of the derivative on the solvability as in the linear systems. In higher dimensions, we have many partial derivatives and we need a systematic procedure to deal with such a complicated case. In other words, we would like to understand the solvability of a system of non-linear equations in several unknowns. This is given via inverse and implicit function theorems. We also remark that in general, it is only possible to obtain a local solvability result and not a global result as in linear systems.

2.1.2 Partial, Directional and Frechét Derivatives

Let $f : \mathbb{R} \to \mathbb{R}$ and $x_0 \in \mathbb{R}$. Then $f'(x_0)$ is normally defined as

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}, \tag{2.5}$$

when the limit exists. We are also aware of the fact that $f'(x_0)$ is the slope of the tangent to the curve $y = f(x)$ at the point $(x_0, f(x_0))$. This allows for another interpretation of the derivative via linear transformation, which is at the heart of the concept of Frechét derivative. Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ be a vector-valued map represented by $f = (f_1, \cdots, f_m)^T$, where $f_i : U \to \mathbb{R}$ are real-valued maps. The limit definition can easily be used to define the directional derivatives in any direction and in particular partial derivatives are nothing but the directional derivatives along the co-ordinate axes.

Directional and Partial Derivatives: Recall that the derivative in (2.5) is the instantaneous rate of change of the output $f(x)$ with respect to the input x. Thus, if we consider $f(x)$ at $x_0 \in \mathbb{R}^n$, there are infinitely many radial directions emanating from x_0. Any given vector $v \in \mathbb{R}^n$ determines a direction given by its position vector. Thus, for $x_0 \in \mathbb{R}^n$, $f(x_0+hv)-f(x_0)$, $h \in \mathbb{R}$ is the change in f in the direction v. This motivates us to define the derivative of f at $x_0 \in \mathbb{R}^n$ in the direction v, denoted by $D_v f(x_0)$, as

$$D_v f(x_0) = \lim_{h\to 0} \frac{f(x_0+hv) - f(x_0)}{h} \tag{2.6}$$

whenever the limit exists. Note that if $f = (f_1, \cdots, f_m)^T$, then

$$D_v f(x_0) = (D_v f_1(x_0), \ldots, D_v f_m(x_0))^T.$$

If v is a unit vector, then $D_v f(x_0)$ is called the *directional derivative* of f at x_0 in the direction v. If $v = e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ is the co-ordinate axis vector, then clearly

$$D_{e_i} f(x_0) = \frac{\partial f}{\partial x_i}(x_0) = \left(\frac{\partial f_1}{\partial x_i}(x_0), \cdots, \frac{\partial f_m}{\partial x_i}\right)^T.$$

Example 2.3. Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = |x|^2$. Then $\dfrac{\partial f}{\partial x_i}(x_0) = 2x_{0i}$. Now, for $v \in \mathbb{R}^n$,

$$\begin{aligned} f(x_0 + hv) &= \sum_{i=1}^{n}(x_{0i} + hv_i)^2 \\ &= f(x_0) + 2h(x_0, v) + h^2|v|^2. \end{aligned}$$

It follows that

$$D_v f(x_0) = 2(x_0, v).$$

□

As seen earlier the existence of all directional derivatives implies the existence of partial derivatives. But, the converse is not true.

Example 2.4. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} x + y & \text{if } x = 0 \text{ or } y = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Then, $D_{(1,0)}f(0, 0) = D_{(0,1)}f(0, 0) = 1$, but $D_{(a,b)}f(0, 0)$, $a \neq 0$, $b \neq 0$ does not exist. □

Normally, we expect differentiable functions to be continuous, which is true in one dimension. But the existence of all directional derivatives at a point does not imply the continuity at that point. This is a serious drawback and prompts us to look for a stronger concept of derivative, namely the notion of *total derivative*.

Example 2.5. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4}, & \text{if } x \neq 0 \\ 0 \text{ if } x = 0. \end{cases}$$

It is easily seen that $D_v f(0, 0)$ exists for all $v \in \mathbb{R}^2$, but f is not continuous at $(0, 0)$. □

Example 2.6. Let the function $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} \dfrac{xy(x^2 - y^2)}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0 \text{ if } (x, y) = (0, 0). \end{cases}$$