Numerical Solution of Partial Differential Equations

Andrea Cangiani[1]
Mathematics Area
SISSA
e-mail: Andrea.Cangiani@sissa.it

January 2021

[1]The bulk of these notes were developed together with Emmanuil Georgoulis at the University of Leicester.

Chapter 1

Introduction to Partial Differential Equations

1.1 Introduction

We give some basic definition, classification, and results from the theory of partial differential equation which are of importance when studying numerical methods for their solution. We refer the reader to any classical book on partial differential equations (eg. Evans') for a more complete presentation.

A partial differential equation (PDE) is an equation involving an unknown function of two or more variables with some of its partial derivatives. PDEs are of fundamental importance in applied mathematics and physics, and have recently shown to be useful in as varied disciplines as financial modelling and modelling of biological systems. More specifically, we have the following definition.

Definition 1.1 *Let $\Omega \subset \mathbb{R}^d$ be an open subset of $\mathbb{R}^d$ (called the* domain of definition*), for $d > 1$ a positive integer (called the* dimension*), and denote by $\mathbf{x} = (x_1, x_2, \dots, x_d)$ a vector in Ω. Let (unknown) function $u : \Omega \to \mathbb{R}$ whose partial derivatives up to order k (for k positive integer) exist. A* partial differential equation of order k in Ω in d dimensions *is an equation of the form:*

$$F(x, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_d}, \dots, \frac{\partial u^k}{\partial x_{d-1} \partial x_d^{k-1}}) = 0, \tag{1.1}$$

where $F : \Omega \times \mathbb{R} \times \mathbb{R}^d \times \cdots \times \mathbb{R}^{d^{k-1}} \times \mathbb{R}^{d^k} \to \mathbb{R}$ is a given function.

Systems of PDEs are analogously defined by considering vector-valued u and F.

We shall be mostly interested in PDEs in two and three dimensions (as these are the ones most often appearing in practical applications), and we shall confine the notation to these cases using (x, y) and (t, x) or (x, y, z) and (t, x, y) to describe two- and three-dimensional vectors respectively (when the notation t is used for an independent variable, this variable should almost always describing "time"). Nevertheless, many properties and ideas described below apply also to the general case of d-dimensions for $d > 3$.

Also, to simplify the notation, we shall often resort to the more compact notation u_x, u_y, u_{xx}, u_{xy}, etc., to signify partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x \partial y}$, etc., respectively.

Definition 1.2 *Consider the notation of Definition 1.1. We call the (classical)* general solution *of the PDE (1.1), the family of functions $u : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ that has continuous partial derivatives up to (and including) order k and that satisfies (1.1).*

There is *no method of solving PDEs that works in general*: different methods work for different families of PDEs. Therefore, it is important to identify such families of PDEs that admit similar properties and, subsequently to describe particular methods of solving PDEs from each such family. Such properties, as we shall see, are also of paramount importance in the selection of the right computational schemes for the numerical solution of the PDE at hand.

1.2 Classification of PDEs

To study PDEs it is often useful to classify them into various families, since PDEs belonging to particular families can be characterised by similar behaviour and properties. There are many and varied classifications for PDEs. Perhaps the most widely accepted and generally useful classification is the distinction between linear and non-linear PDEs. In particular, we have the following definition.

Definition 1.3 *If the PDE (1.1) can be written in the form*

$$a(\mathbf{x})u + b_1(\mathbf{x})u_{x_1} + b_2(\mathbf{x})u_{x_2} + \cdots + b_d(\mathbf{x})u_{x_d} + c_1(\mathbf{x})u_{x_1x_1} + \cdots + c_2(\mathbf{x})u_{x_1x_2} + \cdots + c_{d^2}(\mathbf{x})u_{x_dx_d} + \cdots = f(\mathbf{x}), \tag{1.2}$$

i.e., if the coefficients of the unknown function u and of all its derivatives depend only *on the independent variables $\mathbf{x} = (x_1, x_2, \ldots, x_d)$, then it is called a* linear PDE. *If it is not possible to write (1.1) in the form (1.2), then it is called a* nonlinear PDE.

The family of nonlinear PDEs can be further subdivided into smaller families of PDEs. In particular we have the following definition.

Definition 1.4 *Consider a nonlinear PDE of order k with unknown solution u.*

- *If the coefficients of the k order partial derivatives of u are functions of the independent variables $\mathbf{x} = (x_1, x_2, \ldots, x_d)$* only*, then this is called a* semilinear PDE.
- *If the coefficients of the k order partial derivatives of u are functions of the independent variables $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ and/or of partial derivatives of u of order at most $k-1$ (including u itself) , then this is called a* quasilinear PDE.
- *If a (nonlinear) PDE is not quasilinear, then it is called* fully nonlinear.

Clearly a semilinear PDE is also a quasilinear PDE.

LINEARISATION

Example 1.5 *We give some examples of nonlinear PDEs along with their classifications.*

- *The* reaction-diffusion equation
$$u_t = u_{xx} + u^2,$$
is a semilinear PDE.
- *The* inviscid Burgers' equation
$$u_t + uu_x = 0,$$
is a quasilinear PDE and it is not a semilinear PDE.
- *The* Korteweg-de Vries (KdV) equation
$$u_t + uu_x + u_{xxx} = 0,$$
is a semilinear PDE.
- *The* Monge-Ampère equation
$$u_{xx}u_{yy} - (u_{xy})^2 = 0,$$
is a fully nonlinear PDE.

The above classification of PDEs into linear, semilinear, quasilinear, and fully nonlinear is, roughly speaking, a classification of "increasing difficulty" in terms of studying and solving PDEs. Indeed, the mathematical theory of linear PDEs is now well understood. On the other hand, less is known about semilinear PDEs and quasilinear PDEs, and even less about fully nonlinear PDEs.

1.3 First order linear PDEs

We begin our study of linear PDEs with the case of first order linear PDEs. To simplify the discussion ,we shall only consider equations in 2 dimensions, i.e., for $d = 2$; the case of three or more dimensions can be treated in a completely analogous fashion. We begin with an example.

Example 1.6 *We consider the following linear transport PDE in* $\mathbb{R}^2$*:*

$$u_x + u_y = 0. \tag{1.3}$$

To find its general solution, we perform the following transformation of coordinates (also known as change of variables *in Calculus): we consider new variables* $(\xi, \eta) \in \mathbb{R}^2$ *defined by the transformation of coordinates*

$$(x, y) \to (\xi, \eta) \quad \text{, where } \xi(x, y) = x + y \quad \text{and} \quad \eta(x, y) = y - x.$$

We can also calculate the inverse transformation of coordinates

$$(\xi, \eta) \to (x, y),$$

by solving with respect to x *and* y*, obtaining*

$$x = \frac{1}{2}(\xi - \eta) \quad \text{and} \quad y = \frac{1}{2}(\xi + \eta). \tag{1.4}$$

We write the PDE (1.3) in the new coordinates, using the chain rule from Calculus. Setting $v(\xi, \eta) = u(x(\xi, \eta), y(\xi, \eta))$ *we have, respectively:*

$$u_x = v_\xi \xi_x + v_\eta \eta_x, \qquad u_y = v_\xi \xi_y + v_\eta \eta_y,$$

giving

$$u_x = v_\xi - v_\eta, \qquad u_y = v_\xi + v_\eta.$$

Putting these back to the PDE (1.3), we deduce

$$0 = u_x + u_y = v_\xi - v_\eta + v_\xi + v_\eta = 2v_\xi \quad \text{or} \quad v_\xi = 0. \tag{1.5}$$

Integrating this equation with respect to ξ*, we arrive to*

$$v(\xi, \eta) = f(\eta),$$

for any differentiable function of one variable $f : \mathbb{R} \to \mathbb{R}$*. Using now the inverse transformation of coordinates (1.4), we conclude that the general solution of the PDE (1.3) is given by:*

$$u(x, y) = v(\xi(x, y), \eta(x, y)) = f(\eta(x, y)) = f(y - x).$$

The change of variables $(x, y) \to (\xi, \eta)$ is essentially a clockwise rotation of the axes by an angle $\frac{\pi}{4}$. After rotation, the PDE takes the simpler form (1.5), which can be interpreted geometrically as: v is constant with respect to the variable ξ. In other words, the solution u is constant when $y - x = c$, for any constant $c \in \mathbb{R}$. Hence, the straight lines of the form $y = x + c$ "characterise" the solution of the PDE above; such curves are called *characteristic curves* of a PDE, as we shall see below.

Next, we shall incorporate these ideas into the case of the general first order linear PDE. The general form of a 1st order linear PDE in 2 dimensions can be written as:

$$a(x, y)u_x + b(x, y)u_y + c(x, y)u = g(x, y), \quad \text{for} \quad (x, y) \in \Omega \subset \mathbb{R}^2, \tag{1.6}$$

where a, b, c, g are functions of the independent variables x and y only. We also assume that a, b have continuous first partial derivatives, and that they do *not* vanish simultaneously at any point of the domain of definition Ω. Finally ,we assume that the solution u of the PDE (1.6) has continuous first partial derivatives.

Consider a transformation of coordinates of $\mathbb{R}^2$:

$$(x, y) \leftrightarrow (\xi, \eta),$$

with $\xi = \xi(x, y)$ and $\eta = \eta(x, y)$, which is assumed to be smooth (that is, the functions $\xi(x, y)$ and $\eta(x, y)$ have all derivatives with respect to x and y well-defined) and non-singular, i.e., its Jacobian

$$\frac{\partial(\xi, \eta)}{\partial(x, y)} := \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} = \xi_x \eta_y - \xi_y \eta_x \neq 0, \tag{1.7}$$

in Ω; (this requirement ensures that the change of variables is meaningful, in the sense that it is one-to-one and onto). We also denote by $x = x(\xi, \eta)$ and $y = y(\xi, \eta)$ the inverse transformation, as it will be useful below.

We write the PDE (1.6) in the new coordinates, using the chain rule. Setting $v(\xi, \eta) = u(x(\xi, \eta), y(\xi, \eta))$ we have, respectively:

$$u_x = v_\xi \xi_x + v_\eta \eta_x, \qquad u_y = v_\xi \xi_y + v_\eta \eta_y,$$

giving

$$(a\xi_x + b\xi_y)v_\xi + (a\eta_x + b\eta_y)v_\eta + cv = g(x(\xi, \eta), y(\xi, \eta)), \tag{1.8}$$

after substitution into (1.6). To simplify the above equation, we require that the function $\eta(x, y)$ is such that

$$a\eta_x + b\eta_y = 0; \tag{1.9}$$

if this is the case then (1.8) becomes an ordinary differential equation with respect to the independent variable ξ, whose solution can be found by standard separation of variables.

The equation (1.9) is a slightly simpler PDE of first order than the original PDE. To find the required η we are seek to construct curves such that $\eta(x, y) = \mathit{const}$ for any constant; these are called the *characteristic curves* of the PDE (compare this with the straight lines of the example above).

Differentiating this equation with respect to x, we get

$$0 = \frac{\mathrm{d}\,\mathit{const}}{\mathrm{d}x} = \frac{\mathrm{d}\eta(x, y)}{\mathrm{d}x} = \eta_x \frac{\mathrm{d}x}{\mathrm{d}x} + \eta_y \frac{\mathrm{d}y}{\mathrm{d}x} = \eta_x + \eta_y \frac{\mathrm{d}y}{\mathrm{d}x},$$

where in the penultimate equality we made use of the chain rule for functions of two variables; the above equality yields

$$\frac{\eta_x}{\eta_y} = -\frac{\mathrm{d}y}{\mathrm{d}x}, \tag{1.10}$$

assuming, without loss of generality, that $\eta_y \neq 0$ (for otherwise, we argue as above with the rôles of the x and y variables interchanged, and we get necessarily $\eta_x \neq 0$ from hypothesis (1.7)).

Using (1.10) on (1.9), we deduce the *characteristic equation*:

$$-a\frac{\mathrm{d}y}{\mathrm{d}x} + b = 0, \quad \text{or} \quad \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{b}{a}, \tag{1.11}$$

assuming, without loss of generality that $a \neq 0$ near the point (x_0, y_0) (for otherwise, we have that necessarily $b \neq 0$ near the point (x_0, y_0), as a, b cannot vanish simultaneously at any point due to hypothesis, and we can apply the same argument as above with x and y interchanged). Equation (1.11) is an ordinary differential equation of first order that can be solved using standard separation of variables to give a solution $f(x, y) = \mathit{const}$, say. Setting $\eta = f(x, y)$ and ξ to be any function for which (1.7) holds, we can easily see that (1.9) holds also. Therefore, the PDE (1.6) can be written as

$$(a\xi_x + b\xi_y)v_\xi + cv = g(x(\xi, \eta), y(\xi, \eta)), \quad \text{or} \quad v_\xi + \frac{c}{(a\xi_x + b\xi_y)}v = \frac{g(x(\xi, \eta), y(\xi, \eta))}{(a\xi_x + b\xi_y)},$$

which is an ordinary differential equation of first order with respect to ξ and can be solved using the (standard) method of multipliers to find $v(\xi, \eta)$. Using the inverse transformation of coordinates, we can now find the solution $u(x, y)$ from $v(\xi, \eta)$. This is the so-called *method of characteristics* in finding the solution to a first order PDE.

1.3.1 Problems

Problem 1.7 *Show that the characteristic curves for the PDE*

$$yu_x - xu_y + yu = xy \quad \text{for } y \neq 0,$$

are concentric circles centred at the origin. Then, use the method of characteristics to show that the general solution is

$$u(x,y) = v(\xi(x,y), \eta(x,y)) = \xi(x,y) - 1 + f(\eta(x,y)) = x - 1 + f\big(\frac{x^2+y^2}{2}\big)\mathrm{e}^{-x},$$

for any differentiable function f of one variable.

Problem 1.8 *Use the method of characteristics to find the general solution of the first order linear PDE:*

$$u_x - 2u_y = 0.$$

1.4 Second order linear PDEs

Next up, we have linear PDEs of 2nd order. Here, for simplicity, we shall consider only equations in 2 dimensions, i.e., for $d = 2$. The general form of a 2nd order linear PDE in 2 dimensions can be written as:

$$au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + fu = g, \quad \text{for} \quad (x,y) \in \Omega \subset \mathbb{R}^2, \tag{1.12}$$

where a, b, c, d, e, f, g are functions of the independent variables x and y only. We also assume that a, b, c have continuous second partial derivatives, and that they do *not* vanish simultaneously at any point of the domain of definition Ω. Finally, we assume that the solution u of the PDE (1.12) has continuous second partial derivatives. We shall classify PDEs of the form (1.12) in different types, depending on the sign of the *discriminant* defined by

$$\mathcal{D} := b^2 - ac,$$

at each point $(x_0, y_0) \in \Omega$. More specifically, we have the following definition.

Definition 1.9 *Let $\mathcal{D} = b^2 - ac$ be the discriminant of a second order PDE of the form (1.12) in $\Omega \subset \mathbb{R}^2$ and let a point $(x_0, y_0) \in \Omega$.*

- *If $\mathcal{D} > 0$ at the point (x_0, y_0), the PDE is said to be* hyperbolic *at (x_0, y_0).*
- *If $\mathcal{D} = 0$ at the point (x_0, y_0), the PDE is said to be* parabolic *at (x_0, y_0).*
- *If $\mathcal{D} < 0$ at the point (x_0, y_0), the PDE is said to be* elliptic *at (x_0, y_0).*

The equation is said to be hyperbolic, parabolic or elliptic in the domain Ω if it is, respectively, hyperbolic, parabolic or elliptic at all points of Ω.

Example 1.10 *The following are the archetypal examples of linear second order PDEs $\mathbb{R}^2$.*

- *The* wave equation
$$u_{tt} + cu_{xx} = 0,$$
with $c < 0$, is hyperbolic.
- *The* heat *or* diffusion equation
$$u_t + au_{xx} = 0,$$
with $a < 0$, is parabolic.
- *The* Laplace equation
$$\Delta u := u_{xx} + u_{yy} = 0,$$
is elliptic.

Similarly, in 3 dimensions, the wave equation reads $u_{tt} + c\Delta u = 0$ and heat equation reads $u_t + a\Delta u$, with Δ the Laplace operator with respect to the variables x, y. In turns, the Laplace equation reads $\Delta u(x, y, z) = 0$.

The well known character of the Laplace, diffusion, and wave equations reflect the general character of the classes they represent: elliptic equation are generally time-independent, parabolic equations are time-dependent and usually model diffusion phenomena, and hyperbolic equations are also time-dependent but they model transport, wave-like phenomena which are characterised by finite speed of propagation.

Different *physics* can be modelled locally by considering so-called changing type PDEs.

Example 1.11 *The following are examples of equations of changing type.*

- *The* Grušin equation
$$u_{xx} + x^2 u_{yy} = 0,$$
is elliptic in the set $\{(x,y) \in \mathbb{R}^2 : x \neq 0\}$ *and parabolic in the set* $\{(x,y) \in \mathbb{R}^2 : x = 0\}$.
- *The* Tricomi equation
$$y u_{xx} + u_{yy} = 0,$$
is elliptic in the set $\{(x,y) \in \mathbb{R}^2 : y > 0\}$, *parabolic in the set* $\{(x,y) \in \mathbb{R}^2 : y = 0\}$, *and hyperbolic in the set* $\{(x,y) \in \mathbb{R}^2 : y < 0\}$.

The relevance of the above classification stems from the following result, assuring that applying a change of variables will not alter the type of the PDE.

Theorem 1.12 *The sign of the discriminant $\mathcal{D}$ of a second order PDE of the form (1.12) in $\Omega \subset \mathbb{R}^2$ is invariant under smooth non-singular transformations of coordinates (also known as* change of variables*).*

Proof. Consider a transformation of coordinates of $\mathbb{R}^2$:
$$(x,y) \leftrightarrow (\xi, \eta),$$
with $\xi = \xi(x,y)$ and $\eta = \eta(x,y)$, which is assumed to be smooth (that is, the functions $\xi(x,y)$ and $\eta(x,y)$ have all derivatives with respect to x and y well-defined) and non-singular, i.e., its Jacobian
$$\frac{\partial(\xi,\eta)}{\partial(x,y)} := \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} = \xi_x \eta_y - \xi_y \eta_x \neq 0, \tag{1.13}$$
in Ω. We also denote by $x = x(\xi,\eta)$ and $y = y(\xi,\eta)$ the inverse transformation, as it will be useful below.

We write the PDE (1.12) in the new coordinates, using the chain rule. Setting $v(\xi,\eta) = u(x(\xi,\eta), y(\xi,\eta))$ we have, respectively:
$$u_x = v_\xi \xi_x + v_\eta \eta_x, \qquad u_y = v_\xi \xi_y + v_\eta \eta_y, \tag{1.14}$$
giving
$$\begin{aligned} u_{xx} &= v_{\xi\xi}\xi_x^2 + 2v_{\xi\eta}\xi_x\eta_x + v_{\eta\eta}\eta_x^2 + v_\xi \xi_{xx} + v_\eta \eta_{xx}, \\ u_{yy} &= v_{\xi\xi}\xi_y^2 + 2v_{\xi\eta}\xi_y\eta_y + v_{\eta\eta}\eta_y^2 + v_\xi \xi_{yy} + v_\eta \eta_{yy}, \\ u_{xy} &= v_{\xi\xi}\xi_x\xi_y + v_{\xi\eta}(\xi_x\eta_y + \xi_y\eta_x) + v_{\eta\eta}\eta_x\eta_y + v_\xi \xi_{xy} + v_\eta \eta_{xy}. \end{aligned} \tag{1.15}$$

Inserting (1.14) and (1.15) into (1.12), and factorising accordingly, we arrive to
$$A v_{\xi\xi} + 2B v_{\xi\eta} + C v_{\eta\eta} + D v_\xi + E v_\eta + f v = g, \tag{1.16}$$
where the new coefficients A, B, C, D and E are given by
$$\begin{aligned} A &= a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2, \\ B &= a\xi_x\eta_x + b(\xi_x\eta_y + \xi_y\eta_x) + c\xi_y\eta_y, \\ C &= a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2, \\ D &= a\xi_{xx} + 2b\xi_{xy} + c\xi_{yy} + d\xi_x + e\xi_y, \\ E &= a\eta_{xx} + 2b\eta_{xy} + c\eta_{yy} + d\eta_x + e\eta_y. \end{aligned} \tag{1.17}$$
Thus the discriminant of the PDE in new variables (1.16), is given by
$$\begin{aligned} B^2 - AC &= (a\xi_x\eta_x + b(\xi_x\eta_y + \xi_y\eta_x) + c\xi_y\eta_y)^2 - (a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2)(a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2) \\ &= \cdots = (b^2 - ac)(\xi_x\eta_y - \xi_y\eta_x)^2 = (b^2 - ac)\Big(\frac{\partial(\xi,\eta)}{\partial(x,y)}\Big)^2. \end{aligned} \tag{1.18}$$
This means that the discriminant $B^2 - AC$ of (1.16) has always the same sign as the discriminant $b^2 - ac$ of (1.12), as $\frac{\partial(\xi,\eta)}{\partial(x,y)} \neq 0$ from the hypothesis and, therefore, $\big(\frac{\partial(\xi,\eta)}{\partial(x,y)}\big)^2 > 0$. Since the discriminant of the transformed PDE has always the same sign as the one of the original PDE, the type of the PDE remains invariant. □

Let us now consider some special transformations for PDEs of each type. What we shall see is that, given certain transformation, it is possible to write (1.12) locally in much simpler form, the so-called *canonical form*.

Example 1.13 *Consider the wave equation*

$$u_{xx} - u_{yy} = 0,$$

which as we saw before is hyperbolic in $\mathbb{R}^2$*. Let us also consider the transformation of coordinates of* $\mathbb{R}^2$*:*

$$(x, y) \leftrightarrow (\xi, \eta), \quad \text{with} \quad \xi = x + y \quad \text{and} \quad \eta = x - y.$$

It is, of course, smooth as $x + y$ *and* $x - y$ *are infinite times differentiable with respect to* x *and* y*, and it is non-singular. The transformed equation is given by*

$$4v_{\xi\eta} = 0, \quad \text{or} \quad v_{\xi\eta} = 0.$$

(Note that, indeed, this equation is still hyperbolic.) From this canonical form, we can in fact compute the general solution of the wave equation. Indeed, integrating with respect to η*, we arrive to* $v_\xi = h(\xi)$ *for an arbitrary continuously differentiable function* h*. Integrating now the last equality with respect to* ξ*, we deduce* $v = \int^\xi h(s)\mathrm{d}s + G(\eta)$*. If we set* $F(\xi) := \int^\xi h(s)\mathrm{d}s$*, to simplify the notation, we get* $v(\xi, \eta) = F(\xi) + G(\eta)$*, for an arbitrary twice continuously differentiable function* G*, or equivalently*

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

for all twice continuously differentiable functions F *and* G *of one variable: the solution is the sum of the left-travelling function* F *and the right-travelling function* G*. This formula is due to d'Alembert (1717-83).*

We now investigate the following question: is it always possible to find transformations of coordinates that make the general PDE (1.12) "simpler"?

For the general PDE, we employ a geometric argument. We seek functions $\xi(x, y)$ and $\eta(x, y)$ for which we have

$$a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2 = 0 \quad \text{and} \quad a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2 = 0; \tag{1.19}$$

i.e., $A = C = 0$ for the coefficients of the transformed PDE (1.16). The equations (1.19) are PDEs of first order, for which we are now seeking to construct curves such that $\xi(x, y) = const$ for any constant. When (x, y) are points on a curve, i.e, they are such that $\xi(x, y) = const$, they are dependent. Hence, differentiating this equation with respect to x, we get

$$0 = \frac{\mathrm{d}\, const}{\mathrm{d}x} = \frac{\mathrm{d}\xi(x, y)}{\mathrm{d}x} = \xi_x \frac{\mathrm{d}x}{\mathrm{d}x} + \xi_y \frac{\mathrm{d}y}{\mathrm{d}x} = \xi_x + \xi_y \frac{\mathrm{d}y}{\mathrm{d}x},$$

where in the penultimate equality we made use of the chain rule for functions of two variables; the above equality yields

$$\frac{\xi_x}{\xi_y} = -\frac{\mathrm{d}y}{\mathrm{d}x}, \tag{1.20}$$

assuming, without loss of generality, that $\xi_y \neq 0$. Now, we go back to the desired equations (1.19), and we divide the first equation by ξ_y^2 to obtain

$$a\Big(\frac{\xi_x}{\xi_y}\Big)^2 + 2b\frac{\xi_x}{\xi_y} + c = 0,$$

and, using (1.20), we arrive to

$$a\Big(\frac{\mathrm{d}y}{\mathrm{d}x}\Big)^2 - 2b\frac{\mathrm{d}y}{\mathrm{d}x} + c = 0, \tag{1.21}$$

which is called the *characteristic equation* for the PDE (1.12). This is a quadratic equation for $\frac{\mathrm{d}y}{\mathrm{d}x}$, with discriminant $\mathcal{D} = b^2 - ac$! The roots of the characteristic equation are given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{b \pm \sqrt{\mathcal{D}}}{a}. \tag{1.22}$$

Each of the equations above is a first order ordinary differential equation that can be solved using standard separation of variables to give (families of) solutions $f_1(x, y) = const$ and $f_2(x, y) = const$, say. The curves defined by the equations $f_1(x, y) = const$ and $f_2(x, y) = const$ are called the *characteristic curves* of the second order PDE.

Therefore, if the original PDE (1.12) is hyperbolic, i.e., if $\mathcal{D} > 0$, the characteristic equation has two real distinct roots, giving two real distinct characteristics curves for the PDE. If the original PDE (1.12) is parabolic, thereby $\mathcal{D} = 0$, the characteristic equation has one double root, giving one real characteristic curve for the PDE. Finally, if the original PDE (1.12) is elliptic, thereby $\mathcal{D} < 0$, the characteristic equation has no real roots, and therefore the PDE has **no** real characteristic curves, but as we shall see below it has complex characteristic curves. The characteristic curves can be thought as the "natural directions" in which the PDE "communicates information" to different points in its domain of definition Ω. With this statement in mind, it is possible to see that each type of PDE models different phenomena and also admits different properties, rendering the above classification into hyperbolic, parabolic and elliptic PDEs of great importance.

From the above development it is immediate to prove the following theorems characterising the canonical form for those PDEs admitting real characteristic curves, namely those of hyperbolic or parabolic type.

Theorem 1.14 *Let (1.12) be a hyperbolic PDE. Then, for every $(x_0, y_0) \in \Omega$ there exists a transformation of coordinates $(x, y) \leftrightarrow (\xi, \eta)$ in the neighbourhood of (x_0, y_0), such that (1.12) can be written as*

$$v_{\xi\eta} + \cdots = g, \tag{1.23}$$

where "…" are used to signify the terms involving u, u_x, or u_y. This is called the canonical form *of a hyperbolic PDE.*

Theorem 1.15 *Let (1.12) be parabolic PDE. Then, for every $(x_0, y_0) \in \Omega$ there exists a transformation of coordinates $(x, y) \leftrightarrow (\xi, \eta)$ in the neighbourhood of (x_0, y_0), such that (1.12) can be written as*

$$v_{\xi\xi} + \cdots = g, \tag{1.24}$$

where "…" are used to signify the terms involving u, u_x, or u_y. This is called the canonical form *of a parablic PDE.*

These leaves out the elliptic case (when the characteristic equation (1.21) has no real roots). By other means (eg. the theory of analytic functions) it is still possible to prove that also in this case a canonical form always exists.

Theorem 1.16 *Let (1.12) be an elliptic PDE. Then, for every $(x_0, y_0) \in \Omega$ there exists a transformation of coordinates $(x, y) \leftrightarrow (\xi, \eta)$ in the neighbourhood of (x_0, y_0), such that (1.12) can be written as*

$$v_{\xi\xi} + v_{\eta\eta} + \cdots = g, \tag{1.25}$$

where "…" are used to signify the terms involving u, u_x, or u_y. This is called the canonical form *of an elliptic PDE.*

Remark 1.17 *Notice that the whole discussion in this section about linear second order PDEs will still be valid for the case of semilinear second order PDEs too! Indeed, since in second order semilinear PDEs the non-linearities are not present in the coefficients of the second order derivatives, the calculations and the theorems above will still be valid.*

1.4.1 Problems

Problem 1.18 *Consider the PDE:*

$$(1 - M^2)u_{xx} + u_{yy} = 0.$$

(This equation models the potential of the velocity field of a fluid around a planar obstacle; M is called the Mach number.) What is the type of the above second order linear PDE for different values of M? If you know what a "sonic boom" is, can you see a relation to it and the properties of the equation above?

Problem 1.19 *Calculate the characteristic curves of the Tricomi equation*

$$yu_{xx} + u_{yy} = 0,$$

for $y \leq 0$. Show that, when $y < 0$ the Tricomi equation can be written in the canonical form (1.23)

Problem 1.20 *The Black-Scholes equation for a European call option with value $C = C(\tau, s)$ (τ the time variable and s is the asset price), is given by*

$$C_\tau + \frac{\sigma^2}{2} s^2 C_{ss} + rsC_s - rC = 0, \tag{1.26}$$

where r is a positive constant (the interest rate). What type of 2nd order linear PDE is (1.26) and why? Using the following transformation of coordinates of $\mathbb{R}^2$:

$$(\tau, s) \leftrightarrow (t, x), \quad \text{with} \quad \tau = T - \frac{2t}{\sigma^2}, \quad \text{and} \quad s = \mathrm{e}^x,$$

where T is a constant (the final time), show that (1.26) can be transformed into the following PDE in canonical form:

$$v_{xx} + (k-1)v_x - v_t - kv = 0, \tag{1.27}$$

where $v(t,x) := C(\tau(t,x), s(t,x)) = C(T - 2t/\sigma^2, \mathrm{e}^x)$, and $k := 2r/\sigma^2$. Setting now

$$v(t,x) = \mathrm{e}^{\alpha x + \beta t} u(t,x),$$

for some function $u = u(t,x)$, show that the transformed equation (1.27) can be written as

$$u_t - u_{xx} = 0,$$

when

$$\alpha = -\frac{1}{2}(k-1), \quad \text{and} \quad \beta = -\frac{1}{4}(k+1)^2,$$

i.e., the Black-Scholes equation can be transformed into the heat equation!

1.5 The Cauchy problem and well-posedness of PDEs

We have seen, eg. by the method of characteristics, that the general solution of a PDEs contain unknown functions. These take the role of the unknown constants appearing in the general solution of ODEs.

In this section, we shall study some appropriate conditions that will be sufficient to specify the unknown functions and arrive to unique solutions. Hence we shall introduce the notion of *well-posedness* of a PDE problem.

Definition 1.21 *Consider a PDE of the form (1.1), of order k in Ω in d dimensions and let S be a (given) smooth surface on $\mathbb{R}^d$. Let also $n = n(x)$ denote the unit normal vector to the surface S at a point $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in S$. Suppose that on any point $\mathbf{x}$ of the surface S the values of the solution u and of all its directional derivatives up to order $k-1$ in the direction of n are given, i.e., we are given functions $f_0, f_1, \ldots, f_{k-1} : S \to \mathbb{R}$ such that*

$$u(\mathbf{x}) = f_0(\mathbf{x}), \quad \textit{and} \quad \frac{\partial u}{\partial n}(\mathbf{x}) = f_1(\mathbf{x}), \quad \textit{and} \quad \frac{\partial^2 u}{\partial n^2}(\mathbf{x}) = f_2(\mathbf{x}), \ldots, \quad \textit{and} \quad \frac{\partial^{k-1} u}{\partial n^{k-1}}(\mathbf{x}) = f_{k-1}(\mathbf{x}). \tag{1.28}$$

The Cauchy problem *consists of finding the unknown function(s) u that satisfy simultaneously the PDE and the conditions (1.28). The conditions (1.28) are called the* initial conditions *and the given functions $f_0, f_1, \ldots, f_{k-1}$, will be referred to as the* initial data.

Example 1.22 *Consider the Cauchy problem for the 1st order transport equation*

$$\begin{cases} u_x + u_y = 0, \\ u(0, y) = \sin y \quad \textit{on } S = \{(x, y) \in \mathbb{R}^2 : x = 0\}. \end{cases} \tag{1.29}$$

The characteristic curves of the PDE are $y = x + c$, $c \in \mathbb{R}$; notice that S intersects all *of them. From Example 1.6 we also know that the general solution to the PDE is $u(x, y) = f(y - x)$, for all $(x, y) \in \mathbb{R}^2$. Using the initial condition we deduce that a solution to the Cauchy problem is given by $u(x, y) = \sin(y - x)$.*

Example 1.23 *Consider the Cauchy problem for the wave equation*

$$\begin{cases} u_{xx} - u_{yy} = 0, \\ u(x, 0) = \sin x, \quad \textit{and} \quad u_y(x, 0) = 0. \end{cases} \tag{1.30}$$

Here the surface S in Definition 1.21 is implicitly given as $S = \{(x, y) \in \mathbb{R}^2 : y = 0\}$. Imposing the initial conditions to the general solution to the wave equation from Example 1.13 we deduce that a solution to the Cauchy problem is

$$u(x, y) = F(x + y) + G(x - y) = \frac{1}{2}(\sin(x + y) + \sin(x - y)).$$

One question that arises is whether the solutions to the Cauchy problems in the previous examples are unique. A partial answer to this question is given by the celebrated Cauchy-Kovalevskaya Theorem, the proof of which can be found in any standard PDE theory textbook.

Theorem 1.24 (The Cauchy-Kovalevskaya Theorem) *Consider the Cauchy problem from Definition (1.21) for the case of a linear PDE of the form (1.2). Let $\mathbf{x}_0$ be a point of the initial surface S, which is assumed to be analytic. Suppose that S is* not *a characteristic surface at the point $\mathbf{x}_0$. Assume that all the coefficients of the PDE (1.2), the right-hand side f, and all the initial data $f_0, f_1, \ldots, f_{k-1}$ are analytic functions on a neighbourhood of the point $\mathbf{x}_0$. Then the Cauchy problem has a solution u, defined in the neighbourhood of $\mathbf{x}_0$. Moreover, the solution u is analytic in a neighbourhood of $\mathbf{x}_0$ and it is unique in the class of analytic functions.*

Therefore, according to the Cauchy-Kovalevskaya Theorem (under the analyticity assumptions), the Cauchy problem has a solution which is unique in the space of analytic functions. Even if a PDE problem has a unique solution, this does not necessarily mean that the PDE problem is "well behaved". By well-behaved here we understand if the PDE problem changes "slightly" (e.g., by altering "slightly" some coefficient), then also its solution should change only "slightly" also. In other words, "well behaved" is to be understood as follows: "small" changes in the initial data or the PDE itself should *not* result to arbitrarily "large" changes in the behaviour of the solution to the PDE problem.

Definition 1.25 *A PDE problem is* well-posed *if the following 3 properties hold:*

- *the PDE problem has a solution*
- *the solution is unique*
- *the solution depends continuously on the PDE coefficients and the problem data.*

If a PDE problem is not *well-posed, then we say that it is* ill-posed.

The concept of well-posedness is due to Hadamard[1].

Example 1.26 *The Cauchy problem for the wave equation*

$$\begin{cases} u_{xx} - u_{yy} = 0, \\ u(x,0) = f(x), \quad u_y(x,0) = 0, \end{cases}$$

for some known *initial datum f, is an example of a well posed problem. Indeed, working completely analogously to Example 1.23, we can see that a solution to the above problem is given by*

$$u(x,y) = \frac{1}{2}\big(f(x-y) + f(x+y)\big).$$

The proof of uniqueness of solution is more involved and will be omitted (it is based on the so-called energy property of the wave equation).

Finally, to show the continuity of the solution to the initial data, we consider also the Cauchy problem

$$\tilde{u}_{xx} - \tilde{u}_{yy} = 0, \quad \textit{together with the initial conditions} \quad \tilde{u}(x,0) = \tilde{f}(x), \quad \tilde{u}_y(x,0) = 0,$$

i.e., we consider a different initial condition $\tilde{f}$ for the Cauchy problem, giving a new solution $\tilde{u}$. Working as above, we can immediately see that the solution to this Cauchy problem is given by

$$\tilde{u}(x,y) = \frac{1}{2}\big(\tilde{f}(x-y) + \tilde{f}(x+y)\big).$$

Now, we look at the difference of the solutions of the two Cauchy problems above. We have

$$u(x,y)-\tilde{u}(x,y) = \frac{1}{2}\big(f(x-y)+f(x+y)\big)-\frac{1}{2}\big(\tilde{f}(x-y)+\tilde{f}(x+y)\big) = \frac{1}{2}\Big(\big(f(x-y)-\tilde{f}(x-y)\big)+\big(f(x+y)-\tilde{f}(x+y)\big)\Big).$$

Hence if the difference $f(z) - \tilde{f}(z)$ is small for all $z \in \mathbb{R}$, then the difference $u - \tilde{u}$ will also be small! That is the solution depends continuously on the PDE coefficients and the problem data.

In Chapter 2, we shall consider appropriate conditions for each type of linear second order equations (elliptic, parabolic, hyperbolic), that result to well-posed problems.

1.5.1 Problems

Problem 1.27 *Show that the solution of the Cauchy problem for the wave equation*

$$\begin{cases} u_{xx} - u_{yy} = 0, \\ u(x,0) = f(x), \quad \textit{and} \quad u_y(x,0) = g(x), \end{cases}$$

for some known initial datum f and g is given by d'Alembert's formula

$$u(x,y) = \frac{1}{2}(f(x-y) + f(x+y)) + \frac{1}{2}\int_{x-y}^{x+y} g(s)ds.$$

Problem 1.28 *Find the solution to the Cauchy problem for the wave equation*

$$\begin{cases} u_{xx} - u_{yy} = 0, \\ u(x,0) = 0, \quad u_y(x,0) = g(x), \end{cases}$$

for some known initial datum g. Is this problem well-posed or ill-posed? Why?

[1] Jacques Salomon Hadamard (1865 - 1963), French mathematician

Problem 1.29 *This example is due to Hadamard. Consider the Cauchy problem for the Laplace equation*

$$\begin{cases} u_{xx} + u_{yy} = 0, \quad \textit{for } -\frac{\pi}{2} < x < \frac{\pi}{2}, \quad \textit{and} \quad y > 0, \\ u(x,0) = 0, \quad u_y(x,0) = \mathrm{e}^{-\sqrt{n}} \cos(nx), \quad \textit{for } -\frac{\pi}{2} \le x \le \frac{\pi}{2} \\ u(-\pi/2, y) = 0 = u(\pi/2, y), \quad \textit{for } y \ge 0, \end{cases}$$

for every $n = 1, 3, 5, \ldots$. The solution to this problem can be found using the method of separation of variables to be

$$u(x,y) = \frac{\mathrm{e}^{-\sqrt{n}}}{n} \cos(nx) \sinh(ny).$$

Show that this problem is ill-posed by observing that, while the change in the initial condition $u_y(x,0)$ in function of n is exponentially small, the change in the respective solution is exponentially large.