Chapter 1

Differential and Difference Equations

In this chapter we give a brief introduction to PDEs. In Section 1.1 some simple problems that arise in real-life phenomena are derived. (A more detailed derivation of such problems will follow in later chapters.) We show by a number of examples how they may often be seen as continuous analogues of discrete formulations (i.e., based on difference equations). In Section 1.2 we briefly summarize the terminology used to describe various PDEs. Thus concepts like order and linearity are introduced. In Chapter 2 we shall discuss the classification of the various types of PDEs in more detail. Finally, we introduce difference equations and notions like scheme and stencil, which play a role in numerical approximation, in Section 1.3.

1.1 Introduction

Many phenomena in nature may be described mathematically by functions of a small number of independent variables and parameters. In particular, if such a phenomenon is given by a function of spatial position and time, its description gives rise to a wealth of (mathematical) models, which often result in equations, usually containing a large variety of derivatives with respect to these variables. Apart from the spatial variable(s), which are essential for the problems to be considered, the time variable will play a special role. Indeed, many events exhibit gradual or rapid changes as time proceeds. They are said to have an *evolutionary* character and an essential part of their modeling is therefore based on *causality*; i.e., the situation at any time is dependent on the past. As far as (mathematical) modeling leads to PDEs, the latter will be called evolutionary, i.e., involve the time t as a variable. The other type of problems are often referred to as *steady state*. We will give some examples to illustrate this background.

A typical PDE arises if one studies the flow of quantities like density, concentration, heat, etc. If there are no restoring forces, they usually have a tendency to spread out. In particular, one may, e.g., think of particles with higher velocities (or rather energy) colliding with particles with lower velocities. The former are initially rather clustered. The energy will gradually spread out, mainly because the high-velocity particles collide with other ones, thereby transferring some of the energy. This is called *dissipation*. A similar effect can be observed for a material dissolved in a fluid with concentrations varying in space. Brownian motion will gradually spread out the material over the entire domain. This is called *diffusion*.

Example 1.1 Consider a long tube of cross section A filled with water and a dye. Initially the dye is concentrated in the middle. Let $u(x,t)$ denote the concentration or density (mass per unit length) of the dye at position x and time t; then we see that in a small volume $A\Delta x$, positioned between $x - \frac{1}{2}\Delta x$ and $x + \frac{1}{2}\Delta x$ (Figure 1.1), the total amount of dye equals approximately $u(x,t)\Delta x$. Now consider a similar neighbouring volume $A\Delta x$ between $x + \frac{1}{2}\Delta x$ and $x + \frac{3}{2}\Delta x$, with a corresponding dye concentration $u(x + \Delta x, t)$. The mass that flows per unit time through a cross section is called the mass flux. From the physics of solutions it is known that the dye will move from the volume with higher concentration to one with lower concentration such that the mass flux f between the respective volumes is proportional to the difference in concentration between both volumes and is thus given by

$$f\left(x + \frac{1}{2}\Delta x, t\right) = \alpha\left(u\left(x + \frac{1}{2}\Delta x, t\right)\right)\frac{u(x + \Delta x, t) - u(x,t)}{\Delta x},$$

where α, the diffusion coefficient, usually depends on u. This relation is called Fick's law for mass transport by diffusion, which is the analogue of Fourier's law for heat transport by conduction.

As there is a similar flux between the centre volume and its left neighbour, we have a rate of change of total amount of mass in the centre volume equal to the difference between both fluxes given by

$$\frac{\partial}{\partial t}u(x,t)\Delta x = f\left(x + \frac{1}{2}\Delta x, t\right) - f\left(x - \frac{1}{2}\Delta x, t\right).$$

If the diffusion coefficient α is a constant, we have

$$\frac{\partial}{\partial t}u(x,t) = \alpha\frac{u(x + \Delta x, t) - 2u(x,t) + u(x - \Delta x, t)}{\Delta x^2}. \qquad (*)$$

By taking the limits for small volumes (i.e., $\Delta x \to 0$), we find

$$\frac{\partial}{\partial t}u(x,t) = \alpha\frac{\partial^2}{\partial x^2}u(x,t),$$

which is called the one-dimensional *diffusion equation*. As heat conduction satisfies the same equation, it is also called the *heat equation* if u denotes temperature. □

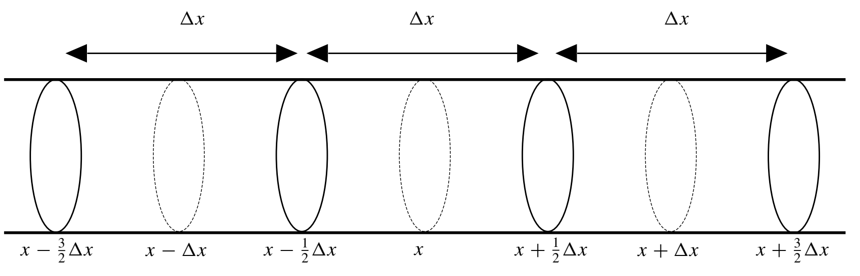

Figure 1.1. *Sketch of dye diffusion.*

Another kind of PDE occurs in the transport of particles. Here a flow typically has a dominant direction; mutual collision of particles (which is felt globally as a kind of internal friction, or viscosity) is neglected.

Example 1.2 Consider a road with heavy traffic moving in one direction, say the x direction (Figure 1.2). Let the number of cars at time t on a stretch $[x, x + \Delta x]$ be denoted by $\Delta N(x, t)$. Furthermore, let the number of cars passing a point x per time period Δt be given by $f(x, t)\Delta t$.

In that period the number of cars $\Delta N(x, t + \Delta t)$ can only be changed by a difference between inflow at x and outflow at $x + \Delta x$; i.e.,

$$\Delta N(x, t + \Delta t) = \Delta N(x, t) - \big(f(x + \Delta x, t) - f(x, t)\big)\Delta t.$$

Rather than the number of cars ΔN per interval of length Δx, it is convenient to consider a *car density* $n(x, t)$, which is defined by

$$\Delta N(x, t) = n(x, t)\Delta x.$$

Hence we obtain the relation

$$\frac{n(x, t + \Delta t) - n(x, t)}{\Delta t} = -\frac{f(x + \Delta x, t) - f(x, t)}{\Delta x}.$$

Assuming sufficient smoothness (which implies that we have to allow for fractions of cars ...), this leads in the limit of $\Delta t, \Delta x \to 0$ to

$$\frac{\partial n}{\partial t} + \frac{\partial f}{\partial x} = 0,$$

which takes the form of a *conservation law*. We may recognize f again as a flux. If this flux only depends on the local car density, i.e., $f = f(n)$, and f is sufficiently smooth, we obtain

$$\frac{\partial n}{\partial t} + f'(n)\frac{\partial n}{\partial x} = 0,$$

also known as the *transport equation*. □

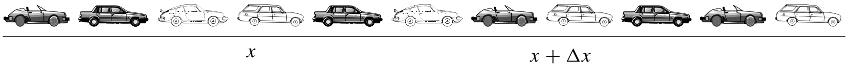

Figure 1.2. *Sketch of traffic flow.*

An important class of problems arises from classical mechanics, i.e., Newtonian systems.

Example 1.3 Consider a chain consisting of elements, each with mass m, and springs, with spring constant $\beta > 0$ and length Δx; see Figure 1.3. Denote the elements by $V_1, V_2, \ldots$ with position of the masses $x = u_1, u_2, \ldots$. Assuming linear springs, the force necessary to increase the original length Δx of the spring of element V_i by an amount $\delta_i = u_i - u_{i-1} - \Delta x$ is equal to $F_i = \beta\delta_i$. Apart from the endpoints, all masses are free to move in the x direction, their inertia being balanced by the reaction forces of the springs. Noting that each element V_i (except for the endpoints) experiences a spring force from the neighbouring ith and $(i + 1)$th springs, we have from Newton's law for the ith element that

$$m\frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} = F_{i+1} - F_i = \beta(u_{i+1} - u_i - u_i + u_{i-1}), \qquad i = 1, 2, \ldots. \qquad (*)$$

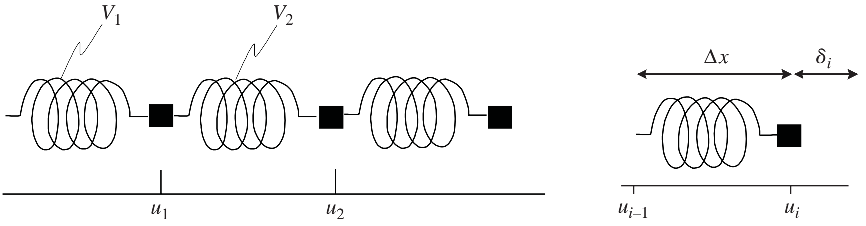

Figure 1.3. *Chain of coupled springs.*

If the chain elements increase in number, while the springs and masses decrease in size, it is natural and indeed more convenient not to distinguish the individual elements, but to blend the discrete description of ($*$) into a continuous analogue. The small masses are conveniently described by a density ρ such that $m = \rho\Delta x$, while the large spring constants are best described by a stiffness $\sigma = \beta\Delta x$. Then we obtain from ($*$) for the position function $u(x, t)$ the PDE

$$\frac{\partial^2 u}{\partial t^2} = \frac{\sigma}{\rho}\frac{\partial^2 u}{\partial x^2}. \qquad (\dagger)$$

As solutions of this equation are typically wave like, it is known as the *wave equation*, with a wave velocity equal to $\sqrt{\sigma/\rho}$. In our example it describes longitudinal waves along the suspended chain of masses. In the context of pressure-density perturbations of a compressible fluid like air, the equation describes one-dimensional sound waves, e.g., as they occur in organ pipes. In that case the air stiffness is equal to $\sigma = \gamma p$, where $\gamma = 1.4$ is a gas constant and p is the atmospheric pressure (see Section 6.8.2). □

In the following example we mention the analogue in electrical circuits of the motion of coupled spring-dashpot elements.

Example 1.4 The time-behaviour of electric currents in a network may be described by the variables potential V, current I, and charge Q. If the network is made of simple wires connecting isolated nodes, resistances, capacities, and coils, and the frequencies are low, it may be modeled (a posteriori confirmed by analysis of the Maxwell equations) one dimensionally by a series of elements with the material properties resistance R, capacitance C, and inductance L. Such a model is called an electrical circuit. If the frequencies are high, such that the wavelength is comparable with the length of the conductors, we have to be more precise. As the signal cannot change instantaneously at all locations, it propagates as a wave of voltage and current along the line. In such a case we cannot neglect the resistance and inductance properties of the wires. By considering the wires as being built up from a series of (infinitesimally) small elements, we can model the system by what is called a transmission line, leading to PDEs in time and space.

In or across each element we have the following relations. The current is defined as the change of charge in time, $I = \frac{\mathrm{d}}{\mathrm{d}t}Q$. The capacitance of a pair of conductors is given by $C = Q/V$, where V is the potential difference and Q is the charge difference between the conductors (Coulomb's law). The resistance between two points is given by $R = V/I$, where V is the potential difference between these points and I is the corresponding current (Ohm's law). A changing electromagnetic current in a coil with inductance L induces a counteracting potential, given by $V = -L\frac{\mathrm{d}}{\mathrm{d}t}I$ (Faraday's law). At a junction no charge can accumulate, and we have the condition $\sum I = 0$, while around a loop the summed potential vanishes. $\sum V = 0$ (Kirchhoff's laws). With these building blocks we can construct transmission line models.

A famous example is the *telegraph equation*, where an infinitesimal piece of telegraph wire is modeled (Figure 1.4) as an electrical circuit consisting of a resistance $R\Delta x$ and an inductance $L\Delta x$, while it is connected to the ground via a resistance $(G\Delta x)^{-1}$ and a capacitance $C\Delta x$.

Let $i(x,t)$ and $u(x,t)$ denote the current and voltage through the wire at position x and time t. The change of voltage across the piece of wire is now given by

$$u(x+\Delta x, t) - u(x,t) = \left[-iR\,\Delta x - \frac{\partial i}{\partial t} L\,\Delta x \right]_{x+\Delta x}.$$

The amount of current that disappears via the ground is

$$i(x+\Delta x, t) - i(x,t) = \left[-uG\,\Delta x - \frac{\partial u}{\partial t} C\,\Delta x \right]_{x}.$$

By taking the limit $\Delta x \to 0$, we get

$$\frac{\partial u}{\partial x} = -Ri - L\frac{\partial i}{\partial t}, \quad \frac{\partial i}{\partial x} = -Gu - C\frac{\partial u}{\partial t}.$$

By eliminating i, we may combine these equations into the telegraph equation for u, i.e.,

$$\frac{\partial^2 u}{\partial x^2} = LC\frac{\partial^2 u}{\partial t^2} + (LG+RC)\frac{\partial u}{\partial t} + RGu. \qquad (*)$$

□

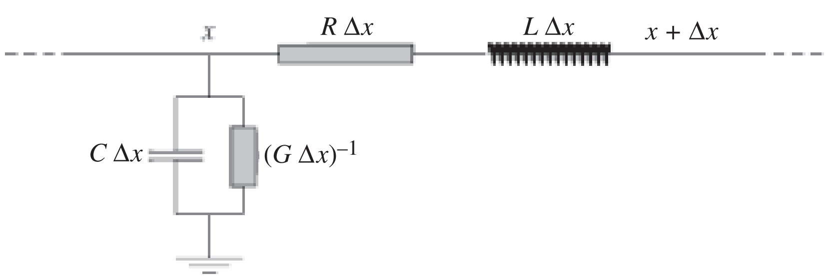

Figure 1.4. *A transmission line model of a telegraph wire.*

Example 1.5 Consider the following crowd of N^2 very accommodating people (Figure 1.5), for convenience ordered in a square of size $L \times L$, while each person, labelled by (i,j), is positioned at $x_i = ih$, $y_j = jh$, with $h = L/N$. Each person has an opinion given by the (scalar) number p_{ij} and can only communicate with his or her immediate neighbours. Assume that each person tries to minimize any conflict with his or her neighbours and is willing to take an opinion that is the average of their opinions. So we have

$$p_{ij} = \frac{1}{4}\left(p_{i+1,j} + p_{i-1,j} + p_{i,j+1} + p_{i,j-1}\right). \qquad (*)$$

Only at the borders of the square are the individuals provided with information such that p is fixed.

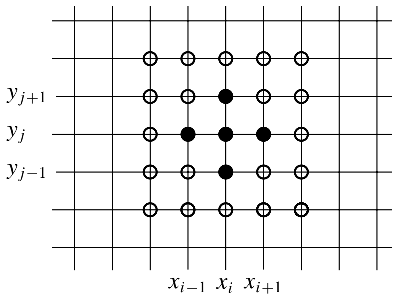

Figure 1.5. *An array of accommodating individuals.*

If the number of people becomes so large that we may take the limit $N \to \infty$ (i.e., $h \to 0$) and p becomes a continuous function of (x, y), $(*)$ becomes

$$p(x, y) = \frac{1}{4}(p(x+h, y) + p(x-h, y) + p(x, y+h) + p(x, y-h)).$$

This may be recast into

$$\big[p(x+h, y) - 2p(x, y) + p(x-h, y)\big] + \big[p(x, y+h) - 2p(x, y) + p(x, y-h)\big] = 0.$$

If this is true for any h, we may divide by h^2, and the equation becomes in the limit

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = 0.$$

This equation is called the *Laplace equation* and describes phenomena where, in some sense, information is exchanged in all directions until equilibrium is achieved. From the above sociological example it is not difficult to appreciate that discontinuities and sharp gradients are smoothed out, while extremes only occur at the boundary. The best-known problem described by this equation is the stationary distribution of the temperature in a heat-conducting medium. □

1.2 Nomenclature

In the previous section we met a number of equations with derivatives with respect to more than one variable. In general, such equations are called *partial differential equations*. Let x and t be two independent variables and let $u(x, t)$ denote a quantity depending on x and t. Furthermore, let

$$t \in [0, T], \quad 0 \le T \le \infty, \quad x \in [a, b] \subset \mathbb{R}. \tag{1.1}$$

For an integer n a general form for a scalar PDE (in two independent variables) reads

$$F\left(\frac{\partial^n u}{\partial t^n}, \frac{\partial^n u}{\partial t \partial x^{n-1}}, \ldots, \frac{\partial^n u}{\partial x^n}, \frac{\partial^{n-1} u}{\partial t^{n-1}}, \ldots, \frac{\partial^{n-1} u}{\partial x^{n-1}}, \ldots, \frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, u, x, t\right) = 0. \tag{1.2}$$

The highest-order derivative is called the *order* of the PDE; not all partial derivatives (except the highest of at least one variable) need to be present. The form (1.2) is an *implicit* formulation, i.e., the highest-order derivative(s), the *principal part*, do(es) not appear explicitly. If the latter is the case, we call it an *explicit* PDE. The generalization to more than two independent variables is obvious.

Example 1.6 Some important examples of PDEs are as follows:

(i) $\dfrac{\partial u}{\partial t} + c\left(1 + \dfrac{3}{2}u\right)\dfrac{\partial u}{\partial x} + \tfrac{1}{6}ch^2\dfrac{\partial^3 u}{\partial x^3} = 0$ (*Korteweg–de Vries equation*).

This is a third order PDE.

(ii) $\dfrac{\partial u}{\partial t} + \dfrac{\partial}{\partial x} f(u) = 0$ (*nonlinear transport equation*).

If f is differentiable, we see that this is a first order PDE in u.

(iii) $\dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} = \varepsilon\dfrac{\partial^2 u}{\partial x^2}$ (*the Burgers' equation*).

If $\varepsilon = 0$, this may be referred to as the inviscid Burgers' equation, which is a special case of the transport equation.

(iv) $\dfrac{\partial^2 u}{\partial t^2} - c^2\dfrac{\partial^2 u}{\partial x^2} - \dfrac{1}{3}h^2\dfrac{\partial^4 u}{\partial x^2 \partial t^2} = 0$ (*linearized Boussinesq equation*).

(v) $EI\dfrac{\partial^4 u}{\partial x^4} - T\dfrac{\partial^2 u}{\partial x^2} + m\dfrac{\partial^2 u}{\partial t^2} = 0$ (*vibrating beam equation*).

(vi) $\dfrac{\partial u}{\partial y}\dfrac{\partial^2 u}{\partial y \partial x} - \dfrac{\partial u}{\partial x}\dfrac{\partial^2 u}{\partial y^2} = \nu\dfrac{\partial^3 u}{\partial y^3}$ (*Prandtl's boundary layer equation*). □

In quite a few cases the order can only be deduced after some (trivial) manipulation.

Example 1.7

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(D(u)\frac{\partial u}{\partial x}\right) = f(x) \qquad (\textit{nonlinear diffusion equation}).$$

It is clear that this PDE is second order. There is no analytical, numerical, or practical need to rework this and have $\frac{\partial^2}{\partial x^2}u$ appear explicitly. □

Usually, the variables are space and/or time. Although the variables in (1.2) are generic, we shall use the symbol t to indicate the *time* variable in general. The variable x will refer to *space*. There are major differences between problems where time does and does not play a role. If the time is not explicitly there, the problem is referred to as a *steady state problem*. If the PDE possesses solutions that evolve explicitly with t, we call it an *evolutionary problem*; i.e., there is *causality*. Most of the theory will be devoted to problems in one space variable. However, occasionally we shall encounter more than one such space variable. Fortunately, problems in more such variables often have many analogues of the one-dimensional case. We shall indicate vectors by boldface characters. So in higher-dimensional space the space variable is denoted by $\boldsymbol{x}$, or by $(x, y, z)^T$. The PDE can still be scalar. We have obvious analogues for vector-dependent variables of the foregoing.

Example 1.8 A few other examples are as follows:

(i) $\dfrac{\partial u}{\partial t} - \alpha\left(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}\right) = 0$ (*heat equation* in three dimensions).

We prefer to write this as $\frac{\partial}{\partial t}u - \alpha\nabla^2 u = 0$. ∇^2 is referred to as the *Laplace operator*.

(ii) $\dfrac{\partial^2 u}{\partial t^2} - c^2\nabla^2 u = 0$ (*wave equation* in three dimensions).

(iii) $\nabla^2 u + k^2 u = 0$ (*Helmholtz or reduced wave equation*).

(iv) $(1 - M^2)\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2} = 0$ (*equation for small perturbations in steady subsonic ($M^2 < 1$) or supersonic ($M^2 > 1$) flow*). □

Sometimes one also denotes a partial derivative of a certain variable by an index:

$$u_t := \frac{\partial u}{\partial t}, \qquad u_{tx} := \frac{\partial^2 u}{\partial t \partial x}. \tag{1.3}$$

If we can write (1.2) as a linear combination of u and its derivatives with respect to x and t, and with coefficients only depending on x and t, the PDE is called *linear*. Moreover, it is called *homogeneous* if it does not depend explicitly on x and/or t. If the PDE is a linear combination of derivatives but the coefficients of the highest derivative, say n, depend on $(n-1)$th order derivatives at most, then we call it *quasi-linear* [29].

For any differential equation we have to prescribe certain initial conditions and boundary conditions for the time and space variable(s), respectively. In evolutionary problems they often both appear as initial boundary conditions. We shall encounter various types and combinations in later chapters.

We finally remark that we may look for solutions that satisfy the PDE in a weak sense. In particular, the derivatives may not exist everywhere on the domain of interest. Again we refer to later chapters for further details.

1.3 Difference Equations

Initially, the actual form of the equations we derived in the examples in Section 1.1 was of a difference equation. Like a PDE, we may define a partial difference equation as any relation between values of $u(x, t)$ where $(x, t) \in \mathcal{F} \subset [a, b] \times [0, T)$, $\mathcal{F}$ being a finite set of points of the domain $[a, b] \times [0, T)$. We shall encounter difference equations when solving a PDE numerically, so they should approximate the PDE in some well-defined way. The simplest way to describe the latter is by defining a *scheme*, i.e., a discrete analogue of the (continuous) PDE. Since we shall mainly deal with finite difference approximations in this book, we perceive a scheme as the result of replacing the differentials by finite differences. To this end we have to indicate some (generic) points in the domain $[a, b] \times [0, T)$ at which the function values $u(x, t)$ are taken. The latter set of points is called a *stencil*. We shall clarify this with some examples.

Example 1.9

(i) Consider Example 1.1 again. If we replace $\frac{\partial}{\partial t}u(x, t)$ in equation $(*)$ by a straightforward discretisation, then we obtain the scheme

$$\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = \alpha \frac{u(x + \Delta x, t) - 2u(x, t) + u(x - \Delta x, t)}{\Delta x^2},$$

and the stencil is the set of bullets (•) in Figure 1.6.

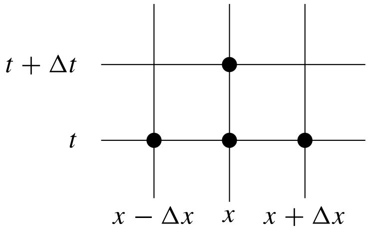

Figure 1.6. *Stencil of Example* 1.9(i).

(ii) Consider the wave equation (†) of Example 1.3. A discrete version may be found to be

$$\frac{u(x,t+\Delta t) - 2u(x,t) + u(x,t-\Delta t)}{\Delta t^2} = \frac{\sigma}{\rho}\frac{u(x+\Delta x,t) - 2u(x,t) + u(x-\Delta x,t)}{\Delta x^2}.$$

The stencil is given in Figure 1.7. □

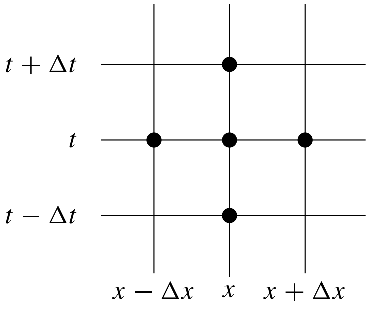

Figure 1.7. *Stencil of Example* 1.9(ii).

Given the special role of time and the implication it has for the actual computation, which should be based on the causality of the problem, we may distinguish schemes according to the number of time levels involved. If $(k+1)$ such time levels are involved, we call the scheme a *k-step scheme*. If the scheme involves only spatial differences at earlier time levels, it is called *explicit*; otherwise it is called *implicit*.

Example 1.10

(i) The schemes in Example 1.9 are both explicit, the first being a one-step and the second a two-step scheme.

(ii) We could also approximate the u_{xx} term in the heat equation at time level $t + \Delta t$ and obtain the scheme

$$\frac{u(x,t+\Delta t) - u(x,t)}{\Delta t} = \alpha\frac{u(x+\Delta x,t+\Delta t) - 2u(x,t+\Delta t) + u(x-\Delta x,t+\Delta t)}{\Delta x^2}.$$

This scheme has the stencil given in Figure 1.8. Clearly, it is an implicit one-step scheme. □

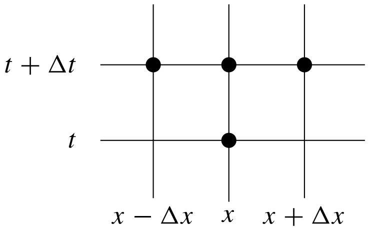

Figure 1.8. *Stencil of Example* 1.10(ii).

1.4 Discussion

- The use of the variables x and y in an equation does not mean that the PDE cannot have an evolutionary character. There are some cases where they refer to spatial coordinates, yet the corresponding equation may be hyperbolic, a type of equation we will encounter in the next chapter as an instance of evolutionary type.
- If in a system of time-dependent PDEs all spatial derivatives are replaced by suitable difference approximations, we obtain a system of ODEs in time. If one of the PDEs is independent of time, we obtain a *differential-algebraic system*. A typical example is the condition that an incompressible flow is divergence free (equivalent to conservation of mass), as in the Stokes equations. This problem will be discussed in Section 8.7.

Exercises

1.1. Show that a nonconstant diffusivity $\alpha(u)$ leads to the equation

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(u)\frac{\partial u}{\partial x}\right).$$

1.2. Determine the order of the *eikonal* equation

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 = c^2.$$

1.3. Determine the order of the PDE

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

Derive a first order system by writing $p := \frac{\partial u}{\partial x},\ q := \frac{\partial u}{\partial y}$.

1.4. Determine the order of the PDE (where a and b are parameters)

$$\frac{\partial u}{\partial t} = a\nabla^2 u + b\frac{\partial u}{\partial x} + c(u).$$

1.5. Verify that the solution $u = u(x,t)$ of the transport equation (cf. Example 1.2 or 1.6(ii))

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0, \qquad u(x,0) = v(x),$$

for sufficiently smooth f is implicitly given by

$$u = v\big(x - f'(u)t\big).$$