1
Mathematical Preliminaries

1.1 Introduction

Partial differential equations emerged when shortcomings with the use of ordinary differential equations were found in the study of vibrations of strings, propagation of sound, waves in liquids and in gravitational attraction. Originally the calculus of partial derivatives was supplied by Euler in a series of papers concerned with hydrodynamics in 1734. This work was extended by D'Alembert in 1744 and 1745 in connection with the study of dynamics.

Partial differential equations are the basis of almost every branch of applied mathematics. Such equations arise from mathematical models of most real life situations. Hence quantum mechanics depends on Schrödinger's equations, fluid mechanics on various forms of Navier–Stokes' equations and electromagnetic theory on Maxwell's equations. Partial differential equations form a very large area of study in mathematics, and are therefore important for both analytical and numerical considerations. The analytical aspects are covered in this text and the numerical aspects in the companion volume, "Numerical methods for partial differential equations".

Inevitably there are many aspects of other branches of mathematics which are pertinent to this work, and the relevant material has been brought together in this chapter to save long digressions later, and to give an element of completeness. The first two sections should be covered at the first reading and form a general introduction to the book as a whole. The later sections deal with a range of related topics that will be needed later, and may be tackled as required.

When the differential equations involve only one independent variable such as $y(t)$ in the equation for simple harmonic motion given by

$$\frac{d^2y}{dt^2} + k^2y = 0 \tag{1.1.1}$$

this is then called an ordinary differential equation. Standard methods are available for the analytic solution of particular classes of such equations such as those with constant coefficients, and these methods are familiar in references such as Nagle and Saff (1993), or the classic, Piaggio (1950). However, it is very easy to write an equation whose closed form solution is not expressible in simple terms such as

$$\frac{d^2y}{dx^2} = xy. \tag{1.1.2}$$

For such a problem the ordinary differential equation itself defines the solution function and is used to derive its analytic properties by such devices as series solutions. Numerical methods come into their own to obtain values of the solution function and again there is a vast literature on this topic which includes Lambert (1990) and Evans (1996).

Partial differential equations follow a similar line, but now the dependent variable is a function of more than one independent variable, and hence the derivatives are all partial derivatives. In view of ordinary differential equations, some types lend themselves to analytic solution, and there is a separate literature on numerical solutions. These aspects form the contents of this book and its companion volume.

The *order* of a partial differential equation is the order of the highest derivative. First-order equations can often be reduced to the solution of ordinary differential equations, which will be seen later in the considerations of characteristics. Second-order equations tend to demonstrate the numerical methods applicable to partial differential equations in general. For the most part, consideration here is limited to linear problems – the nonlinear ones constituting current research problems. Linear problems have the dependent variable and its partial derivatives occurring only to the first degree, hence there are no products of the dependent variable and its derivatives. Hence the equation

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0 \tag{1.1.3}$$

is linear. It is called Laplace's equation, and it will be a major topic in this book, whereas

$$\frac{\partial u}{\partial t} + 6u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0 \tag{1.1.4}$$

is the nonlinear Korteweg–de Vries equation. For solutions of this equation, the method of inverse scattering is employed which is outside the scope of this book, and may be pursued in Ablowitz and Clarkson (1991). A linear equation is said to be *homogeneous* if each term contains either the dependent variable or one of its derivatives, otherwise it is said to be non-homogeneous or inhomogeneous.

The fundamental property of a homogeneous linear problem is that if f_1 and f_2 are solutions then so is $f_1 + f_2$. To begin the discussion, three specific physical applications which prove typical of the problems to be solved are introduced. A classification of second-order equations is then covered, and each of the three physical problems falls into a different class in this categorisation.

The first of these physical problems is the heat or diffusion equation which can be derived by considering an arbitrary volume V. The heat crossing the boundary will equate to the change of heat within the solid, which results in the equation

$$\int_V \rho c \frac{\partial \theta}{\partial t} dV = \int_S k \text{ grad } \theta \cdot d\mathbf{S} + \int_V H(\mathbf{r}, \theta, t)\, dV \tag{1.1.5}$$

where $d\mathbf{S} = \mathbf{n} dS$ with $\mathbf{n}$ the unit outward normal to the surface S of V and dS is a surface element, θ is the temperature, k is the thermal conductivity, ρ the density and c the specific heat. H represents any heat generated in the volume by such action as radioactive decay, electrical heating or chemical action. A short-hand notation, common in continuum mechanics is used here where grad is defined by

$$\text{grad } u = \left\{ \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right\} \tag{1.1.6}$$

which generates a vector from a scalar u. This is often written in the condensed form-grad $u = \nabla u$. A further short-hand notation that will be used where a condensed notation is acceptable is the use of subscripts to denote partial derivatives. Hence the above definition will become

$$\text{grad } u = \{u_x, u_y, u_z\}. \tag{1.1.7}$$

With this definition, the z dependence may be absent in the case of partial differential equations in two independent variables which will be the dominant case in this book. There are two other *vector operators* which are also used in this book, namely div and curl, defined by

$$\text{div}\, \mathbf{a} = \nabla \cdot \mathbf{a} = \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z} \tag{1.1.8}$$

and

$$\text{curl}\, \mathbf{a} = \nabla \times \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_1 & a_2 & a_3 \end{vmatrix} \tag{1.1.9}$$

where $\mathbf{a} = \{a_1, a_2, a_3\}$. Hence the operator div operates on a vector to generate a scalar, and the operator curl operates on a vector and generates a vector. With these definitions in place the derivation of the main equations may now be continued.

The integral over the surface S can be converted to a volume integral by the divergence theorem (Apostol, 1974) to give

$$\int_V \rho c \frac{\partial \theta}{\partial t} dV = \int_V \text{div } (k \text{ grad } \theta) dV + \int_V H(\mathbf{r}, \theta, t)\, dV. \tag{1.1.10}$$

However, this balance is valid for an arbitrary volume and therefore the integrands themselves must satisfy

$$\rho c \frac{\partial \theta}{\partial t} = \text{div } (k \text{ grad } \theta) + H(\mathbf{r}, \theta, t). \tag{1.1.11}$$

In the special but common case in which $k =$ constant, the diffusion equation reduces to

$$\frac{\partial \theta}{\partial t} = K\nabla^2\theta + Q(\mathbf{r}, \theta, t) \tag{1.1.12}$$

with

$$\left.\begin{aligned} K &= \frac{k}{\rho c} \\ Q &= \frac{H}{\rho c} \end{aligned}\right\}.$$

Typical boundary conditions for this equation would include the following.

(i) $\theta(\mathbf{r}, t)$ is a prescribed function of t for every point $\mathbf{r}$ on the boundary surface S.

(ii) The normal flux through the boundary $\frac{\partial \theta}{\partial n}$ is prescribed on S where $\mathbf{n}$ is a normal vector to the surface S.

(iii) The surface radiation is defined over S, for example, by

$$\frac{\partial \theta}{\partial n} = -a(\theta - \theta_0) \tag{1.1.13}$$

which is Newton's law of radiation.

The heat or diffusion equation applies to a very large number of other physical situations. The diffusion of matter such as smoke in the atmosphere, or a dye or pollutant in a liquid is governed by Fick's law

$$\mathbf{J} = -D \text{ grad } c \tag{1.1.14}$$

where D is the coefficient of diffusion and c is the concentration. The vector $\mathbf{J}$ is the diffusion current vector, and therefore c satisfies

$$\frac{\partial c}{\partial t} = \text{div } (D \text{ grad } c) \tag{1.1.15}$$

or

$$\frac{\partial c}{\partial t} = D\nabla^2 c \tag{1.1.16}$$

if D is a constant. Other physical situations, which are modelled by the diffusion equation include neutron slowing, vorticity diffusion and propagation of long electromagnetic waves in a good conductor such as an aerial.

The second of the fundamental physical equations is the wave equation. Consider a small length of a stretched string as shown in Figure 1.1.

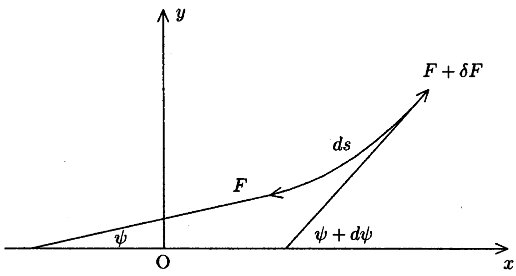

Fig. 1.1.

Then Newton's second law applied to the small length gives

$$F\sin(\psi + d\psi) - F\sin\psi = \rho\, ds\frac{\partial^2 y}{\partial t^2} = F\cos\psi d\psi \tag{1.1.17}$$

to first order where F is the tension in the string, ρ is the string density, ψ is the tangential angle of the string to the x-axis, s is the distance coordinate along the string and y is the displacement from the neutral position. From elementary calculus,

$$\tan\psi = \frac{\partial y}{\partial x} \tag{1.1.18}$$

and hence

$$\sec^2\psi d\psi = \frac{\partial^2 y}{\partial x^2}dx \tag{1.1.19}$$

which yields

$$\begin{aligned} \rho\frac{\partial^2 y}{\partial t^2} &= F\cos^3\psi\frac{\partial^2 y}{\partial x^2}\frac{\partial x}{\partial s} \\ &= F\cos^4\psi\frac{\partial^2 y}{\partial x^2} \end{aligned} \tag{1.1.20}$$

where $\frac{\partial x}{\partial s} = \cos\psi$ as ψ is the angle the tangent makes with the x-axis. However, for oscillations of small amplitude

$$\cos^2\psi = \left\{1 + \left(\frac{\partial y}{\partial x}\right)^2\right\}^{-1} \sim 1 \tag{1.1.21}$$

to yield the wave equation in the form

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} \tag{1.1.22}$$

with

$$c^2 = \frac{F}{\rho}. \tag{1.1.23}$$

The third important example is Laplace's equation which is based on Gauss's law in electromagnetism (see Atkin, 1962) namely

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \rho \tag{1.1.24}$$

where $\mathbf{E}$ is the electric field, and ρ is the charge density with S being the surface of an arbitary volume. If $\rho = 0$, then an application of the divergence theorem, gives the differential equation

$$\text{div } \mathbf{E} = 0. \tag{1.1.25}$$

However, Maxwell's equations with no time varying magnetic field yield

$$\text{curl } \mathbf{E} = 0 \tag{1.1.26}$$

which is the condition for the field to be irrotational. With this proviso, there exists a ϕ such that

$$\mathbf{E} = \text{ grad } \phi \tag{1.1.27}$$

and hence

$$\text{div grad } \phi = 0 \tag{1.1.28}$$

or

$$\nabla^2 \phi = 0 \tag{1.1.29}$$

which is Laplace's equation. The same equation holds for the flow of an ideal fluid. Such a fluid has no viscosity and being incompressible the equation of continuity is div $\mathbf{q} = 0$, where $\mathbf{q}$ is the flow velocity vector (see Acheson, 1990). For irrotational flows the equivalent of 1.1.26 holds to allow the use of the potential function $\mathbf{q} = \text{grad}\phi$ and again the defining equation is 1.1.29.

These three major physical problems (1.1.16, 1.1.21 and 1.1.28) are typical of the main types of second-order linear partial differential equations, and in the next section mathematical arguments will be used to establish a classification of such problems.

The following exercises cover the derivation of variations to the main equations above to allow further physical features to be considered. The mathematical and numerical solutions to these extended problems fall into the remit of the solutions for the basic equations.

EXERCISES

1.1 Establish that if a string is vibrating in the presence of air resistance which is proportional to the string velocity then the wave equation becomes

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} - r\frac{\partial u}{\partial t} \qquad \text{with} \quad r > 0.$$

1.2 Show that if a vibrating string experiences a transverse elastic force (such as a vibrating coiled spring), then the relevant form of the wave equation is

$$\frac{\partial^2 t}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} - ku \qquad \text{with} \quad k > 0.$$

1.3 If a vibrating string is subject to an external force which is defined by $f(x,t)$, then show that the wave equation takes the form

$$\frac{\partial^2 t}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + f(x,t).$$

1.4 If there is an external source or sink of heat given by $f(x,t)$ (such as that due to electrical heating of a rod, chemical heating or radioactive heating), then the diffusion equation becomes

$$\frac{\partial u}{\partial t} = \nabla(k\nabla u) + f(x,t).$$

1.5 If the end of a vibrating string is in a viscous liquid then energy is radiated out from the string to the liquid. Show that the endpoint boundary condition has the form

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

where n is the normal derivative and b is a constant which is positive if energy is radiated to the liquid.

1.6 When a current flows along a cable with leakage, G, the loss of voltage is caused by resistance and inductance. The resistance loss is Ri where R is the resistance and i is the current (Ohm's Law). The inductance loss is proportional to the rate of change of current (Gauss's Law), which gives the term Li_t where L is the inductance. Hence, the voltage equation is

$$v_x + Ri + Li_t = 0.$$

The current change is due to capacitance C, and leakage G. These terms yield

$$i_x + Cv_t + Gv = 0.$$

Deduce the telegraph equation in the form

$$LC\frac{\partial^2 v}{\partial t^2} + (GL + RC)\frac{\partial v}{\partial t} + RGv = \frac{\partial^2 v}{\partial x^2}.$$

Find the equation satisfied by the current i.

1.7 Show that the function

$$u(x,t) = \frac{1}{n} \sin nx \, e^{-n^2 kt}$$

satisfies $u_t = ku_{xx}$. This is a typical separation of the variables solution and will be studied in great detail.

1.2 Characteristics and Classification

Characteristics were first used by Lagrange in two major papers in 1772 and 1779, in which first-order linear and nonlinear equations were considered. This work appears in Chapter 3. Gaspard Monge (1746–1818) introduced characteristic curves in his work "Feuilles d'analyse appliquie à la géométrie" completed in 1770, but not published until 1795, by which time he had considered second-order equations of the type to be discussed next.

Consider a general second-order quasilinear equation defined by the equation

$$Rr + Ss + Tt = W \tag{1.2.1}$$

where

$$R = R(x,y), \qquad S = S(x,y), \qquad T = T(x,y) \qquad \text{and} \qquad W = W(x,y,z,p,q). \tag{1.2.2}$$

with

$$p = \frac{\partial z}{\partial x}, \qquad q = \frac{\partial z}{\partial y}, \qquad r = \frac{\partial^2 z}{\partial x^2}, \qquad s = \frac{\partial^2 z}{\partial x \partial y} \qquad \text{and} \qquad t = \frac{\partial^2 z}{\partial y^2}. \tag{1.2.3}$$

Then the characteristic curves for this equation are defined as curves along which highest partial derivatives are not uniquely defined. In this case these derivatives are the second-order derivatives r, s and t. The set of linear algebraic equations which these derivatives satisfy can be written in terms of differentials, and the condition for this set of linear equations to have a non-unique solution will yield the equations of the characteristics, whose significance will then become more apparent. Hence the linear equations follow by the chain rule as $dz = pdx + qdy$ and also

$$\begin{aligned} dp &= rdx + sdy \\ \text{and} \quad dq &= sdx + tdy \end{aligned} \tag{1.2.4}$$

to give the linear equations

$$\left. \begin{aligned} Rr + Ss + Tt &= W \\ rdx + sdy &= dp \\ sdx + tdy &= dq \end{aligned} \right\} \tag{1.2.5}$$

and there will be no unique solution when

$$\begin{vmatrix} R & S & T \\ dx & dy & 0 \\ 0 & dx & dy \end{vmatrix} = 0 \tag{1.2.6}$$

which expands to give the differential equation

$$R\left(\frac{dy}{dx}\right)^2 - S\left(\frac{dy}{dx}\right) + T = 0. \tag{1.2.7}$$

However, when the determinant in 1.2.6 is zero, we require the linear system 1.2.5 to have a solution. Thus by Cramer's rule,

$$\begin{vmatrix} R & T & W \\ dx & 0 & dp \\ 0 & dy & dq \end{vmatrix} = 0 \tag{1.2.8}$$

also holds, and gives an equation which holds along a characteristic, namely

$$-R\frac{dy}{dx}\, dp - T\, dq + W\, dy = 0. \tag{1.2.9}$$

Returning now to 1.2.7, this equation is a quadratic in dy/dx and there are three possible cases which arise. If the roots are real, then the characteristics form two families of real curves. A partial differential equation 1.2.1 is then said to be of hyperbolic type. The condition is that

$$S^2 - 4RT > 0. \tag{1.2.10}$$

The second case is when

$$S^2 - 4RT = 0 \tag{1.2.11}$$

so the roots are real and equal; in this case we say that equation 1.2.1 is of the parabolic type. When the roots are complex the underlying equation is said to be of elliptic type corresponding to the condition

$$S^2 - 4RT < 0. \tag{1.2.12}$$

The importance of characteristics only becomes apparent at this stage. The first feature is the use of characteristics to classify equations. The methods that will be used subsequently to solve partial differential equations vary from type to type. In the case of hyperbolic equations, the characteristics are real and are used directly in the solution. Characteristics also play a role in reducing equations to a standard or canonical form. Consider the partial differential equation

$$R\frac{\partial^2 z}{\partial x^2} + S\frac{\partial^2 z}{\partial x \partial y} + T\frac{\partial^2 z}{\partial y^2} = G\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) \tag{1.2.13}$$

and put $\xi = \xi(x,y)$, $\eta = \eta(x,y)$ and $z = \zeta$ to see what a general change of variable yields. The result is the partial differential equation

$$A(\xi_x,\xi_y)\frac{\partial^2\zeta}{\partial\xi^2} + 2B(\xi_x,\xi_y,\eta_x,\eta_y)\frac{\partial^2\zeta}{\partial\xi\partial\eta} + A(\eta_x,\eta_y)\frac{\partial^2\zeta}{\partial\eta^2} = F(\xi,\eta,\zeta,\zeta_\zeta,\zeta_\eta) \tag{1.2.14}$$

where

$$A(u,\nu) = Ru^2 + Su\nu + T\nu^2 \tag{1.2.15}$$

and

$$B(u_1,\nu_1,u_2,\nu_2) = Ru_1u_2 + \frac{1}{2}S(u_1\nu_2 + u_2\nu_1) + T\nu_2\nu_2. \tag{1.2.16}$$

The question is now asked for what ξ and η do we get the simplest form? Certainly if ξ and η can be found to make the coefficients A equal to zero, then a simplified form will result. However the condition that A should be zero is a partial differential equation of first order which can be solved analytically (Sneddon (1957) or Vvedensky (1993)). This topic is treated in detail in Chapter 3, and at this stage the reader will need to accept the solution. Different cases arise in the three classifications. In the hyberbolic case when $S^2 - 4RT > 0$, let $R\alpha^2 + S\alpha + T = 0$ have roots λ_1 and λ_2 then $\xi = f_1(x,y)$ and $\eta = f_2(x,y)$ where $f_1(x,y)$ and $f_2(x,y)$ are the solutions of the two factors in the related ordinary differential equations

$$\left[\frac{dy}{dx} + \lambda_1(x,y)\right]\left[\frac{dy}{dx} + \lambda_2(x,y)\right] = 0. \tag{1.2.17}$$

Hence the required transformations are precisely the defining functions of the characteristic curves. With this change of variable, both $A(\xi_x,\xi_y)$ and $A(\eta_x,\eta_y)$ become zero, and the partial differential equation becomes

$$\frac{\partial^2\zeta}{\partial\eta\partial\xi} = \phi(\xi,\eta,\zeta,\zeta_\zeta,\zeta_\eta) \tag{1.2.18}$$

which is the canonical form for the hyperbolic case.

In the parabolic case, $S^2 - 4RT = 0$, there is now only one root, and any independent function is used for the other variable in the transformation. Hence $A(\xi_x,\xi_y) = 0$, but it is easy to show in general that

$$A(\xi_x,\xi_y)A(\eta_x,\eta_y) - B^2(\xi_x,\xi_y,\eta_x,\eta_y) = (4RT - S^2)(\xi_x\eta_y - \xi_y\eta_x)^2$$

and therefore as $S^2 = 4RT$, we must have $B(\xi_x,\xi_y,\eta_x,\eta_y) = 0$ and $A(\eta_x,\eta_y) \neq 0$ as η is an independent function of x and y. Hence when $S^2 = 4RT$, the transformation $\xi = f_1(x,y)$ and $\eta =$ any independent function yields

$$\frac{\partial^2\zeta}{\partial\eta^2} = \phi_1(\xi,\eta,\zeta,\zeta_\zeta,\zeta_n) \tag{1.2.19}$$

which is the canonical form for a parabolic equation. This reduction is shown in the example below.

In the elliptic case there are again two sets of characteristics but they are now complex. Writing $\xi = \alpha + i\beta$ and $\eta = \alpha - i\beta$ gives the real form

$$\frac{\partial^2 \zeta}{\partial \xi \partial \nu} = \frac{1}{4}\left(\frac{\partial^2 \zeta}{\partial \alpha^2} + \frac{\partial^2 \zeta}{\partial \beta^2}\right) \tag{1.2.20}$$

and hence the elliptic canonical form

$$\frac{\partial^2 \zeta}{\partial \alpha^2} + \frac{\partial^2 \zeta}{\partial \beta^2} = \psi(\alpha, \beta, \zeta, \zeta_\alpha, \zeta_\beta). \tag{1.2.21}$$

Note that Laplace's equation is in canonical form as is the heat equation, but the wave equation is not.

In cases where R, S and T are functions of x and y (quasilinear case), then the type of classification will depend on the values of x and y. Hence, there will be regions in the x, y plane in which the conditions for the equation to be parabolic, elliptic and hyperbolic hold. These are regions of parabolicity, ellipticity and hyperbolicity. (see for example Exercise 1.9 below).

As an example of reduction to canonical form consider the quasilinear second-order partial differential equation

$$\frac{\partial^2 u}{\partial x^2} + 2\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} + c^2\frac{\partial u}{\partial y} = 0. \tag{1.2.22}$$

Then the equation of the characteristic curves is

$$\left(\frac{dy}{dx}\right)^2 - 2\frac{dy}{dx} + 1 = 0 \tag{1.2.23}$$

or factorising

$$\left(\frac{dy}{dx} - 1\right)^2 = 0. \tag{1.2.24}$$

The single solution is $x - y =$ const, and therefore the transformation for the canonical form is:

$$\left.\begin{array}{rcl} \xi & = & x - y \\ \eta & = & x \end{array}\right\}. \tag{1.2.25}$$

The choice of η is a conveniently simple arbitrary function which is independent of ξ. The required partial derivatives after this change of variable are

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial \xi^2} + 2\frac{\partial^2 u}{\partial \xi \partial \eta} + \frac{\partial^2 u}{\partial \eta^2} \tag{1.2.26}$$

and

$$\frac{\partial^2 u}{\partial x \partial y} = -\frac{\partial^2 u}{\partial \xi^2} - \frac{\partial^2}{\partial \xi \partial \eta} \tag{1.2.27}$$

which yields the reduced form

$$\frac{\partial^2 u}{\partial x^2} + 2\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial \eta^2}. \tag{1.2.28}$$

Since $\frac{\partial u}{\partial y} = \frac{\partial u}{\partial \xi}$, the transformed equation is

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial \eta^2} = \frac{\partial u}{\partial \xi}. \tag{1.2.29}$$

The reader may now wish to attempt the following exercises on this section.

EXERCISES

1.8 Find the equations of the characteristic curves for the following partial differential equations:

$$\text{(i)} \quad -\frac{\partial^2 u}{\partial x^2} + x\frac{\partial^2 u}{\partial x \partial y} + y^2\frac{\partial^2 u}{\partial y^2} = u + \frac{\partial u}{\partial x}$$

$$\text{(ii)} \quad (1+x^2)\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x \partial y} + (1+y^2)\frac{\partial^2 u}{\partial y^2} = \frac{\partial u}{\partial y}$$

$$\text{(iii)} \quad x^2\frac{\partial^2 u}{\partial x^2} + 2x\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = u$$

and hence deduce which are parabolic, elliptic and hyperbolic.

1.9 Find the regions of parabolicity, ellipticity and hyperbolicity for the partial differential equation:

$$\frac{\partial^2 u}{\partial x^2} + 3xy\frac{\partial^2 u}{\partial x \partial y} + (x+y)\frac{\partial^2 u}{\partial y^2} = u$$

and sketch the resulting regions in the (x, y) plane.

1.10 Find the analytic form of the characteristic curves for the partial differential equation

$$\frac{\partial^2 u}{\partial x^2} + 2\left(x + \frac{1}{y}\right)\frac{\partial^2 u}{\partial x \partial y} + \frac{4x}{y}\frac{\partial^2 u}{\partial y^2} = xy$$

and hence categorise the equation.

1.11 Reduce the equation

$$\frac{\partial^2 z}{\partial x^2} + 2\frac{\partial^2 z}{\partial x \partial y} + 3\frac{\partial^2 z}{\partial y^2} = z$$

to canonical form. Make a further transformation to obtain a real canonical form.

1.3 Orthogonal Functions

There are a number of very important mathematical preliminaries which arise in the solution of partial differential equations, and indeed some of these concepts arise in other branches of mathematics. These preliminaries will be considered in the following sections. The first such topic is that of an orthogonal set of functions. A set of real-valued functions $g_1(x), g_2(x), \ldots$ is called an orthogonal set of functions on an interval (a, b), if they are defined there and if all the integrals $\int_a^b g_n(x)g_m(x)dx$ exist and are zero for all pairs of functions g_n and g_m with $m \neq n$.

The L_2 norm of g_m is

$$||g_m|| = \sqrt{\int_a^b [g_m(x)]^2 dx}. \tag{1.3.1}$$

As an example $\sin\frac{\pi x}{l}, \sin\frac{2\pi x}{l}, \ldots, \sin\frac{n\pi x}{l}, \ldots$ are orthogonal on $0 \leq x \leq l$, since

$$\int_0^l \sin\frac{n\pi x}{l}\sin\frac{m\pi x}{l}dx = 0, \qquad m \neq n \tag{1.3.2}$$

and

$$\left\|\sin\frac{m\pi x}{l}\right\|^2 = \frac{l}{2}. \tag{1.3.3}$$

An orthogonal set of functions $\{g_m\}$ is said to be *complete* if it is impossible to add to the set one other *non-zero* function which is orthogonal to each of the g_m. In other words, $\{g_m\}$ is a complete set if

$$\int_a^b g_m(x)f(x)dx = 0 \tag{1.3.4}$$

for all m implies $f(x) \equiv 0$ on (a, b).

Let $g_1(x), g_2(x), \ldots$ be any complete orthogonal set of functions on $a \leq x \leq b$ and let $f(x)$ be a given function such that

$$f(x) = \sum_{n=1}^{\infty} c_n g_n(x) \tag{1.3.5}$$

where the series converges in the L_2 norm, namely

$$\lim_{m\to\infty}\left\|f - \sum_{n=1}^{m} c_n g_n\right\| \to 0.$$

This is a generalised Fourier series, whose coefficients are found by multiplying both sides by $g_m(x)$ and integrating from a to b to give

$$\begin{aligned}\int_a^b f(x)g_m(x)dx &= \int_a^b \left(\sum_{n=1}^{\infty} c_n g_n(x)\right) g_m(x)dx \\ &= \sum_{n=1}^{\infty} c_n \int_a^b g_n(x)g_m(x)dx \\ &= c_m \int_a^b [g_m(x)]^2\, dx \\ &= c_m||g_m||^2. \end{aligned} \tag{1.3.6}$$

The interchange of the integral with the summation in the second line requires uniform convergence of the series. This happens for certain functions f, the conditions for which can be found in Apostol, (1974 Chapter 11). Hence, the Fourier coefficient c_n is given by

$$c_n = \frac{1}{||g_n||^2} \int_a^b f(x)g_n(x)dx. \tag{1.3.7}$$

For example, if $f(x) = x, \quad 0 < x < l$ and $g_m(x) = \sin \frac{m\pi x}{l}$, then if

$$f(x) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{l} \tag{1.3.8}$$

then

$$c_n = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi x}{l} dx = -\frac{2l}{n\pi}(-1)^n. \tag{1.3.9}$$

Hence the series is

$$f(x) = \sum_{n=1}^{\infty} \frac{2l}{n\pi}(-1)^{n+1} \sin \frac{n\pi x}{l}, \qquad 0 < x < l. \tag{1.3.10}$$

A set of real-valued positive functions $g_1, g_2, \ldots$ is called an orthogonal set of functions with respect to the weight function $\phi(x)$ on the interval (a, b), if they are defined and if all the integrals $\int_a^b \phi(x)g_m(x)g_n(x)dx$ exist and are zero for all pairs of g_m and g_n with $m \neq n$.

The norm of g_m is then

$$||g_m|| = \sqrt{\int_a^b \phi(x)[g_m(x)]^2 dx}. \tag{1.3.11}$$

Similary if $g_1, g_2, \ldots$ is a set of orthogonal functions with respect to a positive weight function ϕ on $a_1 < x < a_2$ and if $f(x) = \sum_{n=1}^{\infty} c_n g_n(x)$, then it can be shown that

$$c_n = \frac{1}{||g_n||^2} \int_a^b f(x)\phi(x)g_n(x)dx. \tag{1.3.12}$$

Consider the set $\{e^{-x} \sin nx : n = 1, 2, 3, \ldots\}$ which is an orthogonal set on $[0, \pi]$ with respect to the weight function e^{2x} since

$$\int_0^{\pi} e^{2x} \left(e^{-x} \sin nx\right) \left(e^{-x} \sin mx\right) dx = 0, \qquad m \neq n. \tag{1.3.13}$$

Moreover

$$||e^{-x} \sin nx||^2 = \int_0^{\pi} e^{2x} \left(e^{-x} \sin nx\right)^2 dx = \frac{\pi}{2}. \tag{1.3.14}$$

In general, most functions met in mathematical physics are sufficiently well-behaved to be expanded using Fourier series in this way, and this allows the wide use of the method of separation of variables which is covered in Chapter 2. A number of exercises on this work may now be attempted.

EXERCISES

1.12 Show that the following sets of functions form orthogonal sets on the stated intervals:

a) $1, \cos x, \cos 2x, \cos 3x, \ldots$ for $0 \leq x \leq \pi$

b) $\sin \pi x, \sin 2\pi x, \sin 3\pi x, \ldots$ for $-1 \leq x \leq 1$

c) $1, 1 - x, 1 - 2x + \frac{1}{2}x^2$ for $0 \leq x \leq \infty$ with respect to the weight function $w(x) = e^{-x}$.

1.13 Show that the functions $f(x) = 1$ and $g(x) = x$ are orthogonal on the interval $-1 < x < 1$. Determine the constants α and β such that the function $h(x) = 1 + \alpha x + \beta x^2$ is orthogonal to both $f(x)$ and $g(x)$.

1.14 Given that $\{1, \cos x, \ldots, \cos nx, \ldots\}$ form a complete set of orthogonal functions on $[0, \pi]$, find $||1||, ||\cos nx||$, and express $x - 2$ in the form $a_0 + \sum_{n=1}^{\infty} a_n \cos nx$ in the interval $[0, \pi]$.

1.15 The Chebyshev polynomials are defined by

$$T_n(x) = \cos(n \arccos x).$$

Show that $T_3(x) = 4x^3 - 3x$ and prove that the orthogonality condition is

$$\int_{-1}^{1} \frac{T_r(x)T_s(x)}{\sqrt{1-x^2}}\,dx = \begin{cases} 0 & r \neq s \\ \pi/2 & r = s \neq 0 \\ \pi & r = s = 0. \end{cases}$$

1.4 Sturm–Liouville Boundary Value Problems

A second area of mathematics which will be treated here in some detail is the theory of Sturm–Liouville problems in ordinary differential equations. In solving

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} \tag{1.4.1}$$

the method of separation of variables is used as described in Chapter 2. This results in ordinary differential equations for two new functions $X(x)$ and $T(t)$. With the equation for X, there is a set of boundary conditions and in solving this system a set of orthogonal functions is found. The method of separation of variables, when applied to second-order linear partial differential equations frequently leads to an ordinary differential equation of the form

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + \{g(x) + \lambda r(x)\}\,\phi(x) = 0 \tag{1.4.2}$$

subject to the boundary conditions

$$\begin{aligned} \alpha_1\phi(a) + \beta_1\phi'(a) &= 0 \\ \alpha_2\phi(b) + \beta_2\phi'(b) &= 0 \end{aligned} \tag{1.4.3}$$

where $p(x), g(x)$ and $r(x)$ are given real continuous functions of x on $a \leq x \leq b$, p is continuously differentiable on (a, b), and such that $p(x) > 0$ (or $p(x) < 0$), $r(x) \geq 0$ (or $r(x) \leq 0$) for $a \leq x \leq b$, ($r(x)$ not identically zero for any part of the range of x considered) and $\alpha_1, \alpha_2, \beta_1, \beta_2$ are constants (at least one non-zero in each boundary condition), λ is a parameter. This system is called a Sturm–Liouville boundary value problem. It generally has non-trivial solutions only for certain λ, these solutions are called eigenfunctions and the corresponding λ are referred to as eigenvalues.

The solution of the linear equation 1.4.2 can be expressed in the form

$$\phi = c_1u_1(x;\lambda) + c_2u_2(x;\lambda) \tag{1.4.4}$$

where u_1 and u_2 are independent solutions and c_1 and c_2 are arbitrary constants.

Substituting into 1.4.3 gives

$$\begin{aligned} c_1[\alpha_1 u_1(a;\lambda) + \beta_1 u_1'(a;\lambda)] + c_2[\alpha_1 u_2(a;\lambda) + \beta_1 u_2'(a;\lambda)] &= 0 \\ c_1[\alpha_2 u_1(b;\lambda) + \beta_2 u_1'(b;\lambda)] + c_2[\alpha_2 u_2(b;\lambda) + \beta_2 u_2'(b;\lambda)] &= 0. \quad (1.4.5) \end{aligned}$$

To ensure that these equations are consistent without implying $c_1 = c_2 = 0$ we must have

$$\begin{vmatrix} \alpha_1 u_1(a;\lambda) + \beta_1 u_1'(a;\lambda) & \alpha_1 u_2(a;\lambda) + \beta_1 u_2'(a;\lambda) \\ \alpha_2 u_1(b;\lambda) + \beta_2 u_1'(b;\lambda) & \alpha_2 u_2(b;\lambda) + \beta_2 u_2'(b;\lambda) \end{vmatrix} = 0. \quad (1.4.6)$$

This equation in λ gives the eigenvalues and the corresponding eigenfunctions are obtained by solving 1.4.2 with this value of λ subject to 1.4.3.

The differential equation may be written $\{L + \lambda r(x)\}y = 0$ where $L = \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + g(x)$. The form L is called the *self-adjoint* form. A bounded operator is self-adjoint if $(Lu, v) = (u, Lv)$ for every u, v in the underlying function space (see Renardy and Rogers, 1993, p253, or Zauderer, 1983, p129). It may appear restrictive, but it is sufficiently general to include most second-order differential operators on $a \le x \le b$. For if

$$L_1 = a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x) \quad (1.4.7)$$

is such an operator then L_1 can be written as

$$\frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + g(x) \quad (1.4.8)$$

by setting

$$p(x) = e^{\int \frac{a_1}{a_2}dx}, \qquad \text{and} \qquad g(x) = \frac{a_0}{a_2}e^{\int \frac{a_1}{a_2}dx}. \quad (1.4.9)$$

Hence for

$$L_1 = x^2\frac{d^2}{dx^2} + \frac{d}{dx} + x^3 \quad (1.4.10)$$

the principle coefficients are

$$a_2 = x^2, \qquad a_1 = 1, \qquad a_0 = x^3$$

and

$$p(x) = e^{\int \frac{1}{x^2}dx} = e^{-\frac{1}{x}}, \qquad g(x) = \frac{x^3}{x^2}e^{-\frac{1}{x}} = xe^{-\frac{1}{x}}. \quad (1.4.11)$$

Therefore L_1 in self-adjoint form is

$$\frac{d}{dx}\left(e^{-\frac{1}{x}}\frac{d}{dx}\right) + xe^{-\frac{1}{x}}. \quad (1.4.12)$$

Two important theorems relating to Sturm–Liouville problems are as follows.

Theorem 1.1

If the Sturm–Liouville problem

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + \{g(x) + \lambda r(x)\}\, y = 0 \tag{1.4.13}$$

with

$$\alpha_1 y(a) + \beta_1 y'(a) = 0, \qquad \alpha_2 y(b) + \beta_2 y'(b) = 0 \tag{1.4.14}$$

has p, q and r real valued and continuous on $a \le x \le b$, p continuously differentiable on (a, b) and has $r(x)$ positive (or negative) everywhere in the whole interval then all the eigenvalues of the problem are real.

Theorem 1.2

The eigenfunctions that correspond to distinct eigenvalues of the Sturm–Liouville system are orthogonal with respect to the weight $r(x)$.

Proofs of these theorems can be found in Lomen and Mark (1988) or Burkill (1962).

1.5 Legendre Polynomials

The analytic solution of partial differential equations will result in the use of many transcendental functions which may be new to the reader. The main group of these form orthogonal sets on various intervals. Some of these functions will be considered here and the objective is to derive the usual relations which appear in compendia such as Abramowitz and Stegun (1964) such as recurrence relations, differential equations, generating functions, special values, formulae for derivatives, series and asymptotic forms.

Given a non-negative integer n, Legendre's equation for the Legendre polynomial of degree n is

$$(1 - x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0, \qquad x \in (-1, 1). \tag{1.5.1}$$

This equation can be written in the form

$$x^2\frac{d^2y}{dx^2} + xp(x)\frac{dy}{dx} + q(x)y = 0 \tag{1.5.2}$$

where

$$p(x) = \frac{-2x^2}{1 - x^2}, \qquad q(x) = \frac{n(n+1)x^2}{1 - x^2}. \tag{1.5.3}$$

These functions can be expanded as convergent series' in powers of x for $|x| < 1$, hence the solution y to 1.5.2 can itself be expressed as a convergent series in x, for x in the interval $-1 < x < 1$, and has the form:

$$P_n(x) = \sum_{m=0}^{[n/2]} \frac{(-1)^m(2n-2m)!x^{n-2m}}{2^n m!(n-m)!(n-2m)!} \tag{1.5.4}$$

where

$$[n/2] = \begin{cases} n/2 & n \text{ even} \\ (n-1)/2 & n \text{ odd.} \end{cases} \tag{1.5.5}$$

The polynomial $P_n(x)$ is called the Legendre polynomial of degree n.

The generating function for the Legendre polynomials has the form

$$(1-2tx+t^2)^{-1/2} = \sum_{t=0}^{\infty} t^l P_l(x) \tag{1.5.6}$$

with $|t| < 1$ and $|x| < 1$.

Proof

Expand $(1-2tx+t^2)^{-1/2}$ by the binomial theorem to yield

$$\begin{aligned}(1-2tx+t^2)^{-1/2} &= [1-t(2x-t)]^{-1/2} \\ &= \sum_{r=0}^{\infty} \frac{(2r)!}{2^{2r}(r!)^2} t^r (2x-t)^r. \end{aligned} \tag{1.5.7}$$

Now expanding $(2x-t)^r$ by the binomial theorem gives

$$(2x-t)^r = \sum_{p=0}^{r} \frac{r!(2x)^{r-p}(-t)^p}{p!(r-p)!} \tag{1.5.8}$$

which implies

$$(1-2tx+t^2)^{-1/2} = \sum_{r=0}^{\infty} \frac{(2r)!}{2^{2r}(r!)^2} \sum_{p=0}^{r} \frac{r!}{p!(r-p)!} (-1)^p t^{r+p} (2x)^{r-p}. \tag{1.5.9}$$

The coefficients of t^l are required, hence let $r+p=l$, and for fixed r we have $p = l-r$; now $0 \le p \le r$, so we must only consider r such that $0 \le l-r \le r$ or $l/2 \le r \le l$. Hence if l is even, r can take values between $l/2$ and l, while if l is odd r can take values between $(l+1)/2$ and l. For any r in these ranges the coefficient of t^l is obtained by taking $p = l-r$ to give

$$\frac{(2r)!r!(-1)^{l-r}(2x)^{2r-l}}{2^{2r}(r!)^2(l-r)!(2r-l)!} \tag{1.5.10}$$

and the total coefficient of t^l is

$$\sum_{r=\alpha}^{l} \frac{(2r)!r!(-1)^{l-r}(2x)^{2r-l}}{2^{2r}(r!)^2(l-r)!(2r-l)!} = \sum_{k=0}^{\beta}\left(\frac{(2l-2k)!(l-k)!(-1)^k(2x)^{l-2k}}{2^{2l-2k}[(l-k)!]^2 k!(l-2k)!}\right) \tag{1.5.11}$$

where $k = l - r$ and

$$\alpha = \begin{cases} l/2 & l \text{ even} \\ (l+1)/2 & l \text{ odd} \end{cases} \qquad \beta = \begin{cases} l/2 & l \text{ even} \\ (l-1)/2 & l \text{ odd} \end{cases}$$

$$= \sum_{k=0}^{[l/2]} \frac{(-1)^k(2l-2k)!x^{l-2k}}{2^l(l-k)!(l-2k)!k!} = P_l(x). \tag{1.5.12}$$

Some explicit expressions for Legendre polynomials which follow from the series 1.5.4 above are:

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{1}{2}(3x^2 - 1),$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x), \qquad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3). \tag{1.5.13}$$

A commonly used relationship to form orthogonal functions is Rodrigues' formulae

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l. \tag{1.5.14}$$

There are a number of special properties for Legendre Polynomials which arise from the definitions and these include

$$\text{(i) } P_l(1) = 1, \qquad \text{(ii) } P_l(-1) = (-1)^l, \qquad \text{(iii) } P_l'(1) = \frac{1}{2}l(l+1),$$

$$\text{(iv) } P_l'(-1) = (-1)^{l-1}\frac{l}{2}(l+1), \qquad \text{(v) } P_{2l}(0) = \frac{(-1)^l(2l)!}{2^{2l}(l!)^2},$$

$$\text{(vi) } P_{2l+1}(0) = 0 \tag{1.5.15}$$

and the derivatives of the Legendre functions satisfy recurrence relations of the form

$$\begin{array}{ll} \text{(i)} & (2l+1)xP_l(x) = (l+1)P_{l+1}(x) + lP_{l-1}(x) \\ \text{(ii)} & P_l(x) = P_{l+1}'(x) - 2xP_l'(x) + P_{l-1}'(x) \\ \text{(iii)} & P_{l+1}'(x) - P_{l-1}'(x) = (2l+1)P_l(x) \\ \text{(iv)} & xP_l'(x) - P_{l-1}'(x) = lP_l(x) \\ \text{(v)} & P_l'(x) - xP_{l-1}'(x) = lP_{l-1}(x) \\ \text{(vi)} & (x^2-1)P_l'(x) = lxP_l(x) - lP_{l-1}(x). \end{array} \tag{1.5.16}$$

Functions that are defined and are square integrable on the interval $(-1, 1)$ can be expanded as series of Legendre Polynomials. Assume that $f(x)$ can be expressed in the interval $-1 \leq x \leq 1$ as an infinite series of Legendre polynomials of the form

$$f(x) = \sum_{n=0}^{\infty} A_n P_n(x). \tag{1.5.17}$$

The generalised Fourier coefficients A_n are required. If 1.5.17 is multiplied by $P_n(x)$ and integrated w.r.t. x from -1 to 1, the orthogonality of the set P_n on $[-1, 1]$, namely

$$\int_{-1}^{1} P_n(x) P_m(x)\, dx = 0 \qquad (m \neq n), \tag{1.5.18}$$

may be utilised. This follows since $P_n(x)$ satisfies Legendre's equation

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_n(x)}{dx}\right] + n(n+1)P_n(x) = 0 \tag{1.5.19}$$

and multiplying through by $P_m(x)$ and integrating w.r.t. x from -1 to 1 gives

$$\int_{-1}^{1} P_m(x)\frac{d}{dx}\left[(1-x^2)\frac{dP_n(x)}{dx}\right] dx + n(n+1)\int_{-1}^{1} P_m(x)P_n(x)\, dx = 0. \tag{1.5.20}$$

Integrate the first term by parts and since $(1-x^2) = 0$ when $x = \pm 1$, we obtain

$$-\int_{-1}^{1} (1-x^2)\frac{dP_n(x)}{dx}\frac{dP_m(x)}{dx}\, dx + n(n+1)\int_{-1}^{1} P_m(x)P_n(x)\, dx = 0. \tag{1.5.21}$$

On the other hand, starting with $P_m(x)$, and multipying through by $P_n(x)$ would give

$$-\int_{-1}^{1} (1-x^2)\frac{dP_n(x)}{dx}\frac{dP_m(x)}{dx}\, dx + m(m+1)\int_{-1}^{1} P_m(x)P_n(x)\, dx = 0. \tag{1.5.22}$$

Subtracting 1.5.21 from 1.5.22 gives

$$[n(n+1) - m(m+1)]\int_{-1}^{1} P_m(x)P_n(x)\, dx = 0 \tag{1.5.23}$$

and if $m \neq n$ we have

$$\int_{-1}^{1} P_m(x)P_n(x) = 0. \tag{1.5.24}$$

Hence the set $\{P_n\}$ is orthogonal on $[-1, 1]$.

If $m = 0$ we have $P_0(x) = 1$ which gives

$$\int_{-1}^{1} P_n(x) = 0 \qquad \text{if } n \neq 0. \tag{1.5.25}$$

Since the highest power of x with a non-zero coefficient occurring in $P_n(x)$ is x^n, any polynomial of degree n in x, $Q_n(x)$ say, can be represented by

$$Q_n(x) = \sum_{r=0}^{n} B_r P_r(x) \tag{1.5.26}$$

where the B_r's are constants. Therefore

$$\int_{-1}^{1} P_m(x)Q_n(x)\,dx = \sum_{r=0}^{n} B_r \int_{-1}^{1} P_m(x)P_r(x)\,dx = 0 \qquad \text{if } m > n. \tag{1.5.27}$$

To evaluate

$$\int_{-1}^{1} [P_n(x)]^2\,dx \tag{1.5.28}$$

we make use of relation (iii) from 1.5.16, multiply through by $P_n(x)$ and integrate from -1 to 1 to obtain

$$(2n+1)\int_{-1}^{1} [P_n(x)]^2\,dx = \int_{-1}^{1} P'_{n+1}(x)P_n(x)\,dx - \int_{-1}^{1} P'_{n-1}(x)P_n(x)\,dx. \tag{1.5.29}$$

However, $P'_{n-1}(x)$ is a polynomial of degree $n-2$ in x, so the last integral is zero. Integrate first by parts to give

$$(2n+1)\int_{-1}^{1} [P_n(x)]^2 dx = P_{n+1}(x)P_n(x)|_{-1}^{1} - \int_{-1}^{1} P_{n+1}(x)P'_n(x)dx. \tag{1.5.30}$$

However, $P'_n(x)$ is a polynomial of degree $n-1$ in x and the last integral is zero, which leaves

$$(2n+1)\int_{-1}^{1} [P_n(x)]^2\,dx = 1 - (-1)^{n+1}(-1)^n = 2. \tag{1.5.31}$$

Hence

$$\int_{-1}^{1} [P_n(x)]^2\,dx = \frac{2}{2n+1}. \tag{1.5.32}$$

Multiplying

$$f(x) = \sum_{n=0}^{\infty} A_n P_n(x) \tag{1.5.33}$$

by $P_m(x)$ and integrating from -1 to 1 gives

$$\begin{aligned} \int_{-1}^{1} f(x) P_m(x)\,dx &= \sum_{n=0}^{\infty} A_n \int_{-1}^{1} P_n(x) P_m(x)\,dx \\ &= A_m \int_{-1}^{1} [P_m(x)]^2\,dx = A_m \frac{2}{2m+1} \end{aligned} \tag{1.5.34}$$

and the required coefficient is

$$A_m = \frac{2m+1}{2} \int_{-1}^{1} f(x) P_m(x)\,dx. \tag{1.5.35}$$

The following exercises illustrate the above work and form a link with the section on orthogonal functions.

EXERCISES

1.16 Show that

$$\int_{-1}^{1} x P_l(x) P_{l-1}(x)\,dx = \frac{2l}{4l^2 - 1}.$$

1.17 Show that

$$\int_{-1}^{1} (1 - x^2) P_l'(x) P_m'(x)\,dx = \frac{2l(l+1)}{2l+1} \delta_{ml}$$

where

$$\delta_{ml} = \begin{cases} 1 & m = l \\ 0 & m \neq l. \end{cases}$$

1.18 If

$$f(x) = \begin{cases} \frac{1}{2} & 0 < x < 1 \\ -\frac{1}{2} & -1 < x < 0 \end{cases}$$

expand $f(x)$ in the form

$$\sum_{r=0}^{\infty} c_r P_r(x).$$

1.19 If

$$u_n = \int_{-1}^{1} x^{-1} P_n(x) P_{n-1}(x)\, dx$$

show that

$$nu_n + (n-1)u_{n-1} = 2$$

and hence show that

$$u_n = \begin{cases} \frac{2}{n} & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

1.20 Show that

$$(1-x)\sum_{r=0}^{n}(2r+1)P_r(x) = (n+1)[P_n(x) - P_{n+1}(x)].$$

1.21 Show that

$$\sum_{r=0}^{n}(2r+1)P_r(x) = P'_{n+1}(x) + P'_n(x).$$

1.22 If n is a positive integer prove that

$$\int_{-1}^{1} P_n(x)(1-2xh+h^2)^{-1/2} dx = \frac{2h^n}{2n+1}.$$

1.23 Verify relations (iii)–(vi) of 1.5.16.

1.6 Bessel Functions

The second of our special functions to be considered is the Bessel function. The generating function for Bessel functions of integer order is

$$\exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right] = \sum_{n=-\infty}^{\infty} t^n J_n(x) \tag{1.6.1}$$

where $J_n(x)$ denotes the Bessel function of the first kind of order n. We expand

$$\exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right] \tag{1.6.2}$$

in powers of t to give:

$$\begin{aligned}
\exp\left[\frac{x}{2}\left(t-\frac{1}{t}\right)\right] &= \exp\left(\frac{xt}{2}\right)\exp\left(-\frac{x}{2t}\right) \\
&= \sum_{r=0}^{\infty}\frac{\left(\frac{1}{2}xt\right)^r}{r!}\sum_{s=0}^{\infty}\frac{\left(-\frac{1}{2}\frac{x}{t}\right)^s}{s!} \\
&= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{(-1)^s}{r!s!}\left(\frac{x}{2}\right)^{r+s}t^{r-s}. \qquad (1.6.3)
\end{aligned}$$

Now pick out the coefficients of t^n, $n \geq 0$. For a fixed value of r we want $s = r - n$, and for this value we have

$$\frac{(-1)^{r-n}}{r!(r-n)!}\left(\frac{x}{2}\right)^{2r-n}. \qquad (1.6.4)$$

Total coefficients of t^n are obtained by summing over all values of r. Since $s = r - n$ and we require $s \geq 0$ we must have $r \geq n$. Hence we have

$$\sum_{r=n}^{\infty}\frac{(-1)^{r-n}}{r!(r-n)!}\left(\frac{x}{2}\right)^{2r-n} = \sum_{p=0}^{\infty}\frac{(-1)^p\left(\frac{x}{2}\right)^{2p+n}}{(p+n)!p!} = J_n(x). \qquad (1.6.5)$$

If $n < 0$, we still have the coefficient of t^n for fixed r given by

$$\frac{(-1)^{r-n}}{r!(r-n)!}\left(\frac{x}{2}\right)^{2r-n} \qquad (1.6.6)$$

but now the requirement $s \geq 0$ with $s = r - n$ is satisfied by all r.

The coefficient of t^n is

$$\sum_{r=0}^{\infty}\frac{(-1)^{r-n}}{r!(r-n)!}\left(\frac{x}{2}\right)^{2r-n} = J_n(x) \qquad (1.6.7)$$

for n negative.

Writing n positive in 1.6.7 gives

$$\begin{aligned}
J_{-n}(x) &= \sum_{r=0}^{\infty}\frac{(-1)^{r+n}\left(\frac{x}{2}\right)^{2r+n}}{r!(r+n)!} \\
&= (-1)^n\sum_{r=0}^{\infty}\frac{(-1)^r\left(\frac{x}{2}\right)^{2r+n}}{r!(r+n)!} \\
&= (-1)^n J_n(x). \qquad (1.6.8)
\end{aligned}$$

Hence

$$J_n(x) = \sum_{p=0}^{\infty}\frac{(-1)^p\left(\frac{x}{2}\right)^{2p+n}}{p!(p+n)!} \qquad (1.6.9)$$

and

$$J_{-n}(x) = (-1)^n J_n(x) \tag{1.6.10}$$

(n positive integer).

An integral representation for $J_n(x)$ is given by

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(n\phi - x\sin\phi)\, d\phi \tag{1.6.11}$$

for integer n, the proof of which can be found in Watson (1922).

The next set of formulae to be found are the three-term recurrence relations:

$$\text{(i)} \quad J_n(x) = \frac{x}{2n}[J_{n-1}(x) + J_{n+1}(x)] \tag{1.6.12}$$

$$\text{(ii)} \quad J_n'(x) = \frac{1}{2}[J_{n-1}(x) - J_{n+1}(x)]. \tag{1.6.13}$$

These formulae can be derived from the generating function 1.6.1.

Differentiate w.r.t. t to obtain

$$\exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right]\frac{x}{2}\left(1 + \frac{1}{t^2}\right) = \sum_{n=-\infty}^{\infty} nt^{n-1}J_n(x) \tag{1.6.14}$$

which gives

$$\frac{x}{2}\sum_{n=-\infty}^{\infty} t^n J_n(x) + \frac{x}{2}\sum_{n=-\infty}^{\infty} t^{n-2}J_n(x) = \sum_{n=-\infty}^{\infty} nt^{n-1}J_n(x). \tag{1.6.15}$$

The coefficient of t^{n-1} yields

$$J_n(x) = \frac{x}{2n}\left[J_{n-1}(x) + J_{n+1}(x)\right]. \tag{1.6.16}$$

Differentiating 1.6.1 w.r.t. x gives

$$\exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right]\frac{1}{2}\left(t - \frac{1}{t}\right) = \sum_{n=-\infty}^{\infty} nt^{n-1}J_n'(x) \tag{1.6.17}$$

which then results in

$$\frac{1}{2}\sum_{n=-\infty}^{\infty} t^{n+1}J_n(x) - \frac{1}{2}\sum_{n=-\infty}^{\infty} t^{n-1}J_n(x) = \sum_{n=-\infty}^{\infty} t^n J_n'(x). \tag{1.6.18}$$

The coefficient t^n gives

$$J_n'(x) = \frac{1}{2}\left[J_{n-1}(x) - J_{n+1}(x)\right] \tag{1.6.19}$$

Eliminating $J_{n+1}(x)$ from 1.6.16 and 1.6.19 gives

$$xJ_{n-1}(x) = nJ_n(x) + xJ_n'(x) \tag{1.6.20}$$

multiply by x^{n-1} to deduce

$$x^n J_{n-1}(x) = nx^{n-1}J_n(x) + x^n J_n'(x) \tag{1.6.21}$$

which gives

$$x^n J_{n-1}(x) = \frac{d}{dx}(x^n J_n(x)). \tag{1.6.22}$$

Similarly eliminating $J_{n-1}(x)$ gives

$$xJ_{n+1}(x) = nJ_n(x) - xJ_n'(x) \tag{1.6.23}$$

multiply by x^{-n-1} to get

$$x^{-n} J_{n+1}(x) = -\frac{d}{dx}(x^{-n} J_n(x)). \tag{1.6.24}$$

From 1.6.22 and 1.6.24,

$$x^{n+1} J_n(x) = \frac{d}{dx}\left[x^{n+1} J_{n+1}(x)\right] = -\frac{d}{dx}\left[x^{2n+1}\frac{d}{dx}[x^{-n} J_n(x)]\right] \tag{1.6.25}$$

which yields

$$x^2 J_n''(x) + xJ_n'(x) + (x^2 - n^2)J_n(x) = 0. \tag{1.6.26}$$

Hence $y = J_n(x)$ is a solution of the differential equation

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2)y = 0 \tag{1.6.27}$$

which is called Bessel's equation for the Bessel function of order n.

Bessel's equation of order ν where ν is not necessarily an integer is

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - \nu^2)y = 0 \tag{1.6.28}$$

by the obvious extension of 1.7.27. This equation has a regular singular point at 0 and an irregular singular point at ∞, and a series solution gives

$$J_\nu(x) = \sum_{s=0}^{\infty} \frac{(-1)^s}{\Gamma(\nu+1+s)s!}\left(\frac{x}{2}\right)^{2s+\nu}. \tag{1.6.29}$$

This is Bessel's function of the first kind of order ν. If ν is an integer n we obtain the series as before. If ν is not an integer we can construct a series based on $c = -\nu$ giving a general solution

$$y = aJ_\nu(x) + bJ_{-\nu}(x) \tag{1.6.30}$$

where

$$J_{-\nu}(x) = \sum_{s=0}^{\infty} \frac{(-1)^s}{\Gamma(-\nu+1+s)s!} \left(\frac{x}{2}\right)^{2s-\nu}. \tag{1.6.31}$$

If ν is a positive integer or zero, $J_{-\nu}(x)$ is a multiple of $J_\nu(x)$. An independent second solution can be found that involves a log term.

Note that $J_0(0) = 1$, $J_n(0) = 0$ for $n \neq 0$ and integer. There are other types of Bessel functions such as spherical Bessel functions and modified Bessel functions. The relationships that exist between them are too numerous to mention and can be found in Abramowitz and Stegun (1964).

The reader is now in a position to consider the following exercises in which further properties of the Bessel functions are developed.

EXERCISES

Show that

1.24

$$J_{\frac{1}{2}}(x) = \left(\frac{2}{\pi x}\right)^{\frac{1}{2}} \sin x$$

1.25

$$J_{\frac{-1}{2}}(x) = \left(\frac{2}{\pi x}\right)^{\frac{1}{2}} \cos x$$

1.26

$$\frac{d}{dx}\left[x J_n(x) J_{n+1}(x)\right] = x\left[J_n^2(x) - J_{n+3}^2(x)\right]$$

1.27

$$8J_n'''(x) = J_{n-3}(x) - 3J_{n-1}(x) + 3J_{n+1}(x) - J_{n+1}(x)$$

1.28

$$4J_0'''(x) + 3J_0'(x) + J_3(x) = 0$$

1.29

$$J_n(-x) = (-1)^n J_n(x).$$

1.30 Using the generating function prove that

$$1 = J_0^2(x) + 2J_1^2(x) + 2J_2^2(x) + \dots.$$

Note that

$$1 = \exp\left[\frac{x}{2}\left(t - \frac{1}{t}\right)\right] \exp\left[-\frac{x}{2}\left(t - \frac{1}{t}\right)\right].$$

1.31 By proving

$$\frac{d}{dx}\left[\frac{x^2}{2}(J_n^2(x) - J_{n-1}(x)J_{n+1}(x))\right] = xJ_n^2(x)$$

show that

$$\int_0^x tJ_n^2(t)dt = \frac{x^2}{2}(J_n^2(x) - J_{n-1}(x)J_{n+1}(x)).$$

1.32 Given that

$$\int_0^\infty J_n(x)dx = 1$$

by using

$$\frac{x}{2}[J_{n+1}(x) + J_{n-1}(x)] = nJ_n(x)$$

show that

$$\int_0^\infty \frac{J_n(x)}{x}dx = \frac{1}{n}.$$

1.33 Show that

$$\begin{aligned}\left(\frac{1}{x}\frac{d}{dx}\right)^r [x^n J_n(x)] &= x^{n-r}J_{n-r}(x)\\ \left(\frac{1}{x}\frac{d}{dx}\right)^r [x^{-n} J_n(x)] &= (-1)^r x^{-n-r}J_{n+r}(x).\end{aligned}$$

1.7 Results from Complex Analysis

Complex analysis is employed quite freely in certain sections of this book, in particular in Chapter 4 on transform methods. The intention here is to quote the most important theorems and explain some of the practical consequences which are used in this text. For a full account of this subject, the reader should refer to a standard text such as Copson (1935), Titchmarsh (1932), Needham (1997) or Stewart and Tall (1983).

A complex number z is an ordered pair, namely its real part and its imaginary part. This is conventionally written as

$$z = x + iy. \tag{1.7.1}$$

The modulus of a complex number z is written as $|z|$ and defined as $\sqrt{x^2+y^2}$. A complex number can be represented by a point in the (x, y) plane (called an Argand diagram), and hence can also be represented in polar form with

$$x = r\cos\theta \qquad \text{and} \qquad y = r\sin\theta. \tag{1.7.2}$$

The angle θ is called the argument of z. The complex conjugate of z is denoted by $\bar{z}$ and defined by

$$\bar{z} = x - iy. \tag{1.7.3}$$

The concepts of continuity and differentiability follow by analogy with the real case. A function which is one-valued and differentiable at every point of a domain of the Argand plane, except at a finite number of points, is said to be analytic in that domain. The exceptional points are called singularities. An analytic function with no singularities is said to be regular.

Some familiar complex functions then include:

$$\exp(z) = 1 + \sum_{n=1}^{\infty} \frac{z^n}{n!} \tag{1.7.4}$$

$$\sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \tag{1.7.5}$$

$$\cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}, \tag{1.7.6}$$

from which it follows that

$$\sin(z) = (e^{iz} - e^{-iz})/2i \tag{1.7.7}$$

$$\cos(z) = (e^{iz} + e^{-iz})/2 \tag{1.7.8}$$

and the commonly used result that

$$e^{i\theta} = \cos(\theta) + i\sin(\theta). \tag{1.7.9}$$

The hyperbolic functions can be defined in the usual way as

$$\sinh z = \frac{1}{2}(e^z - e^{-z}) \tag{1.7.10}$$

$$\cosh z = \frac{1}{2}(e^z + e^{-z}) \tag{1.7.11}$$

to yield the relationships

$$\begin{aligned} \sin iz &= i\sinh z & \cos iz &= \cosh z \\ \sinh iz &= i\sin z & \cosh iz &= \cos z. \end{aligned} \tag{1.7.12}$$

As with real numbers, the logarithmic function can be defined as the solution $\log z$ of the equation $e^w = z$. There are an infinite number of solutions in the complex case. These have the form

$$\log z = \log |z| + i \arg z \tag{1.7.13}$$

and the multiplicity of solutions is realised by $\arg z$ having many values each differing from the next by 2π.

One of the most far-reaching theorems in complex analysis is Cauchy's theorem. If $z = x(t) + iy(t)$ then in the Argand plane z describes a curve as t varies. If the arc is replaced by n small straight lines segments, then the arc is said to be rectifiable if the sum of the lengths of these segments tends to a unique limit as $n \to \infty$. Hence the integration of a complex function along an arc in the Argand plane is reduced to the integration of two real functions. Cauchy's theorem states that if z_1 and z_2 are two points in a complex domain D joined by a rectifiable arc lying in D then the value of the integral from z_1 to z_2 is quite independent of the particular arc employed, provided the integrand is analytic in the region considered. The theorem is usually expressed in the form that if $f(z)$ is an analytic function, continuous within and on a simple closed rectifiable curve C, then if $f'(z)$ exists at each point within C, then

$$\int_C f(z)\, dz = 0. \tag{1.7.14}$$

A useful consequence of Cauchy's Theorem is Cauchy's integral theorem which states that if $f(z)$ is analytic, and is regular within a closed contour C (without self-intersections) and is continuous within and on C, then for any point a within C:

$$f(a) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z-a}\, dz \tag{1.7.15}$$

where C is traversed in an anticlockwise sense. As $f(z)$ is regular at a then

$$f(z) = f(a) + (z-a)f'(a) + (z-a)\eta(z) \tag{1.7.16}$$

where $\eta \to 0$ as $z \to a$. Hence by Cauchy's theorem, if γ is a circle centred at a with radius $r < \delta$ where $|z-a| < \delta$ then

$$\begin{aligned} \int_C \frac{f(z)}{z-a}\, dz &= \int_\gamma \frac{f(z)}{z-a}\, dz \\ &= f(a) \int_\gamma \frac{dz}{z-a} + f'(a) \int_\gamma dz + \int_\gamma \eta\, dz \\ &= 2\pi i f(a) + \int_\gamma \eta\, dz. \end{aligned} \tag{1.7.17}$$

However $\left| \int_\gamma \eta\, dz \right| \le 2\pi r \max_{|z-a|<r} |\eta(z)|$, and hence the final integral tends to zero with r to yield the theorem. By a similar means it can be shown that

$$f'(a) = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z-a)^2}\, dz \tag{1.7.18}$$

and more generally

$$f^{(n)}(a) = \frac{n!}{2\pi i}\int_C \frac{f(z)}{(z-a)^{n+1}}\,dz \tag{1.7.19}$$

where again C is traversed in the anticlockwise sense.

A consequence of these results is Taylor's theorem which states that if $f(z)$ is an analytic function regular in a neighbourhood of $z = a$, then $f(z)$ may be expanded in the form

$$f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n. \tag{1.7.20}$$

and the series converges for all z s.t. $|z-a| < R$ for some $R > 0$. Consider a point $a + h$, then Cauchy's integral theorem gives

$$\begin{aligned} f(a+h) &= \frac{1}{2\pi i}\int_C \frac{f(z)}{z-a-h}\,dz \\ &= \frac{1}{2\pi i}\int_C f(z)\left\{\frac{1}{z-a} + \frac{h}{(z-a)^2} + \dots \right. \\ &\quad \left. + \frac{h^n}{(z-a)^{n+1}} + \frac{h^{n+1}}{(z-a)^{n+1}(z-a-h)}\right\} dz. \end{aligned} \tag{1.7.21}$$

Now use the derivative expressions in 1.7.20 to yield

$$f(a+h) = f(a) + \sum_{r=1}^{n} f^{(r)}(a)\frac{h^r}{r!} + \text{remainder}. \tag{1.7.22}$$

If a function is not regular in the domain $|z-a| < R_1$ but is regular in the annulus $R_2 < |z-a| < R_1$ then the expansion has the form

$$f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n + \sum_{n=1}^{\infty} b_n(z-a)^{-n} \tag{1.7.23}$$

which is Laurent's theorem. If at $z = a$, the Laurent expansion is a terminating series in $1/(z-a)^n$ then the singularity is called a pole of order m if b_m is the last non-zero coefficient. The coefficient b_1 is called the residue of $f(z)$ at the pole a. This value has considerable importance in evaluating complex integrals. For a simple pole of order unity the residue is given by

$$b_1 = \lim_{z\to a}(z-a)f(z) \tag{1.7.24}$$

and the formula for a pole of order m is given in the exercises.

Being able to evaluate residues is important in the evaluation of contour integrals, and is called the calculus of residues, the important theorem being Cauchy's theorem of residues. This theorem states that if $f(z)$ is continuous within and on a closed contour C, and if $f(z)$ is regular except for a finite number of poles within C, then

$$\int_C f(z)\,dz = 2\pi i \text{ (sum of the residues of } f(z) \text{ at its poles within } C\text{).} \tag{1.7.25}$$

This result is easily seen by considering one specific pole of order m say at a. Then $f(z)$ has the form

$$f(z) = g(z) + \sum_{n=1}^{m} \frac{a_n}{(z-a)^n}, \tag{1.7.26}$$

where g is a regular function in the neighbourhood of a. The contour C may be deformed by Cauchy's theorem to a circle defined by $z = a + \epsilon e^{i\theta}$ centred on the pole. Then

$$\int_C f(z)\,dz = \sum_{n=1}^{m} a_n \epsilon^{1-n} \int_0^{2\pi} e^{(1-n)i\theta} i\,d\theta = 2\pi i a_1 \tag{1.7.27}$$

as required. For a set of poles, the contour C may be distorted round each pole to be summed to give the theorem.

As an example consider the integral

$$\int_0^\infty \frac{dx}{1+x^2} = \frac{\pi}{2}. \tag{1.7.28}$$

To evaluate this integral using the calculus of residues consider

$$\int \frac{dz}{1+z^2} \tag{1.7.29}$$

around a contour which is defined by the real axis from $-R$ to R closed by the semicircle in the upper half plane. The integrand has one pole inside this contour at $z = i$, and its residue is $1/2i$, making the contour integral have value π. However, the contour integral is equal to twice the required integral plus the integral round the semicircle as $R \to \infty$. Putting $z = Re^{i\theta}$ gives for the integral round the semicircle

$$\int_0^\pi \frac{Rie^{i\theta}}{1+R^2e^{2i\theta}}\,d\theta \tag{1.7.30}$$

which tends to zero as $R \to \infty$. Some exercises are now presented to ensure familiarity with these ideas.

EXERCISES

1.34 Prove the the residue b_1 for pole of order m at $z = a$ is given by

$$b_1 = \lim_{z\to a} \frac{d^{m-1}}{dz^{m-1}}(z-a)^m f(z).$$

1.35 Show by contour integration that

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$$

by considering the contour integral

$$\int_C \frac{e^{iz}}{z}\,dz$$

over a contour consisting of the real axis from r to R, a semicircle centre the origin radius R in the upper half plane, the real axis from $-R$ to $-r$ and a small semicircle radius r back to $z = r$. Let $R \to \infty$ and $r \to 0$.

1.36 Use contour integration to evaluate the integral

$$\int_0^\infty \frac{\cos x}{a^2 + x^2}\,dx.$$

1.8 Generalised Functions and the Delta Function

This section discusses the background to generalised functions which is used extensively in Chapter 5 on Green's functions. The reader is advised to consider this work as a prelude to Chapter 5, rather than consider this material at the first reading.

Since the mid 1930s, engineers and physicists have found it convenient to introduce fictitious functions having idealised properties that no physically significant functions can possibly possess. The main reason for this was their use in solving engineering problems often compounded in terms of a partial differential equation. For example, if a partial differential equation is taken to represent some dynamical engineering system, it is often useful to know how this system behaves when it is disturbed by an impulse. The delta function is a commonly used example of a generalised function, and can be thought of as an idealised mathematical representation of an impulse. The solution to the partial differential equation using a δ-impulse can therefore provide a model for the behaviour of the system.

The Dirac delta function is defined by some authors as the function having the properties

$$\delta(x) = \begin{cases} 0, & x \neq 0; \\ \infty, & x = 0; \end{cases}$$

$$\int\limits_{-\infty}^{\infty} f(x)\delta(x)\,dx = f(0). \qquad (1.8.1)$$

This function was first introduced by PAM Dirac in the 1930s in order to facilitate the analysis of partial differential equations in quantum mechanics; although the idea had been around for a century or more in mathematical circles. Clearly, such a function does not exist in the sense of classical analysis. Dirac called it an "improper function" and its use in analysis is recommended only when it is obvious that no inconsistency will follow from it (self-consistency being one of the principal foundations of any mathematical and scientific discipline). It is interesting to note that Dirac was also the first to postulate the existence of antimatter; like the delta function, another very abstract concept considered to be rather absurd at the time.

When attempting to provide a rigorous interpretation of the above equations, it is necessary to generalise the concept of a function. It was the work of L Schwartz and MJ Lighthill in the 1950s which put the theory of $\delta(x)$, and another fictitious functions, on a firm foundation. The mathematical apparatus developed by Schwartz and Lighthill is known as the "Theory of distributions". Today, there exist other approaches which put the Dirac delta function on a firm basis including "non-standard analysis".

1.8.1 Definition and Properties of a Generalised Function

To obtain an idea of what a generalised function is, it is convenient to use as an analogy, the notion of an irrational number ξ being a sequence $\{\rho(n)\}$ of rational numbers $\rho(n)$ such that

$$\xi = \lim_{n\to\infty} \rho(n),$$

where the limit indicates that the points $\rho(n)$ on the real line converge to the point representing ξ. All arithmetic operations performed on the irrational number ξ are actually performed on the sequence $\{\rho(n)\}$ defining ξ. A generalised function can be thought of as being a sequence of functions, which when multiplied by a test function and integrated over the whole sequence yields a finite limit. This approach is analogous to the one developed by K. Weierstrass in the 18th century who expressed a differential in terms of the limit of a variable approaching zero (but not actually reaching zero), and gives

$$\frac{df}{dx} = \lim_{\delta x\to 0} \frac{f(x+\delta x) - f(x)}{\delta x}.$$

This limiting argument avoided the issue associated with the fact that when $\delta x = 0$ we have 0/0! – one of the principal criticisms of calculus until the concept of a limit was introduced. A list of definitions and results from the theory of distributions is now given.

(i) Test Functions

The definition of test functions is as follows, where **R** and **C** are taken to denote a set of real and complex numbers respectively. The term "iff" denotes the statement "if and only if".

A function $\phi: \mathbf{R} \to \mathbf{C}$ is said to be a test function iff:

(i) $\phi \in C^\infty(\mathbf{R}, \mathbf{C})$ (i.e. ϕ is infinitely differentiable);

(ii) $|x^i \phi^{(j)}(x)| \leq M_{ij}$, for all integers $i, j \geq 0$ and all x in **R**.

Here, $C^\infty(\mathbf{R}, \mathbf{C})$ denotes the linear space of all **C**-valued and continuous functions defined on **R** such that their derivatives, $\phi^{(j)}$, of all orders are continuous.

The set of test functions is denoted by $S(\mathbf{R}, \mathbf{C})$. However, for certain generalised functions, this class can be extended, e.g. for the delta function it is sufficient to assume that the test functions are continuous.

An example of a function belonging to $S(\mathbf{R}, \mathbf{C})$ is the Gaussian function

$$\text{gauss}(x) = \frac{1}{\sqrt{\pi}} \exp(-x^2). \tag{1.8.2}$$

(ii) Regular Sequences

A sequence $\{\phi(x;n)\} \subset S(\mathbf{R}, \mathbf{C})$ is said to be regular iff the following limit exists for any $f \in S(\mathbf{R}, \mathbf{C})$:

$$\lim_{n\to\infty} \langle \phi(x;n), f(x) \rangle \equiv \lim_{n\to\infty} \int_{-\infty}^{\infty} \phi(x;n) f(x)\, dx. \tag{1.8.3}$$

For this limit to exist, it is not necessary that the sequence converges pointwise. For example, the sequence $\{n\,\text{gauss}(nx)\}$ approaches infinity as $n \to \infty$ at the point $x = 0$. However, the above limit exists. Also, even when the pointwise limit of a regular sequence does exist everywhere, it does not need to have any connection with the above limit. For example, the sequence $\{\exp(-1/n^2x^2)/n^2x^2\}$ approaches zero everywhere (except at $x = 0$ where the entries of the sequence are undefined), as $n \to \infty$, whereas the above limit approaches to $f(0)$.

(iii) Equivalent Regular Sequences

Two regular sequences $\{\phi(x;n)\}, \{\psi(x;n)\}$ are said to be equivalent iff for all $f \in S(\mathbf{R}, \mathbf{C})$

$$\lim_{n\to\infty} \langle \phi(x;n), f(x) \rangle = \lim_{n\to\infty} \langle \psi(x;n), f(x) \rangle. \tag{1.8.4}$$

For example, the sequences $\{n\,\text{gauss}(nx)\}$ and $\{2n\,\text{gauss}(2nx)\}$ are equivalent, leading to the limit $f(0)$. Generalised functions are now defined in terms of equivalent regular sequences.

(iv) A Generalised Function

ψ is a generalised function iff ψ is defined as the total, or complete class of equivalent regular sequences.

The term total in the above definition means that there exists no other equivalent regular sequences not belonging to this class. Any member of the class, for example $\{\psi(x;n)\}$, is sufficient to represent both ψ and the total class of equivalent regular sequences defining ψ. This is symbolically denoted as $\psi \sim \{\psi(x;n)\}$.

(v) The Functional $\langle\psi,\phi\rangle$

The functional $\langle\psi,\phi\rangle$ is defined as

$$\langle\psi,\phi\rangle \equiv \lim_{n\to\infty} \int_{-\infty}^{\infty} \psi(x;n)\phi(x)\,dx \equiv \int_{-\infty}^{\infty} \psi(x)\phi(x)\,dx. \tag{1.8.5}$$

where ψ is a generalised function and $\phi \in S(\mathbf{R},\mathbf{C})$.

The last integrand on the right-hand side of the above definition is used symbolically and does not imply actual integration. For some purposes it is also convenient to introduce the set whose elements are called functions of slow growth.

(vi) Functions of Slow Growth

$f\colon \mathbf{R} \to \mathbf{C}$ is said to be of slow growth iff

(i) $f \in C^{\infty}(\mathbf{R},\mathbf{C})$;

(ii) for each $j, j = 0,1,2,\ldots$, there exists a $B_j > 0$ such that $|f^{(j)}(x)| = O(|x|^{B_j})$, as $|x| \to \infty$.

The set of functions of slow growth will be denoted as $N(\mathbf{R},\mathbf{C})$. From this definition it is clear that any polynomial is an element of $N(\mathbf{R},\mathbf{C})$. Moreover, if $a \in N(\mathbf{R},\mathbf{C})$ and $\phi \in S(\mathbf{R},\mathbf{C})$, then $a\phi \in S(\mathbf{R},\mathbf{C})$. The elements of $S(\mathbf{R},\mathbf{C})$ are known as good functions and those of $N(\mathbf{R},\mathbf{C})$ as fairly good functions.

The algebraic operations for generalised functions are now defined. In this definition, ϕ and ψ are generalised functions represented by the sequences

$\phi(x) \sim \{\phi(x;n)\}$ and $\psi(x) \sim \{\psi(x;n)\}$ respectively.

(vii) Algebra of Generalised Functions

1 Addition: $\phi(x) + \psi(x) \sim \{\phi(x;n) + \psi(x;n)\}$.

2 Multiplication by a scalar: $\alpha\phi(x) \sim \{\alpha\phi(x;n)\}$, $\alpha \in \mathbf{C}$.

3 Derivative: $\phi'(x) \sim \{\phi'(x;n)\}$.

4 Shifting similarity: $\phi(\alpha x + \beta) \sim \{\phi(\alpha x + \beta; n)\}$, $\alpha, \beta \in \mathbf{C}$.

5 Multiplication by elements of $N(\mathbf{R}, \mathbf{C})$: $a(x)\phi(x) \sim \{a(x)\phi(x;n)\}$.

Note that the operation of multiplication between two generalised functions is not defined in general. From the above definition the following properties can be derived which are presented without proof.

(viii) Properties of Generalised Functions

If $\phi \in S(\mathbf{R}, \mathbf{C})$, ψ is a generalised function and $a(x) \in N(\mathbf{R}, \mathbf{C})$, then

1 $\langle \psi^{(l)}, f \rangle = (-1)^l \langle \psi, f^{(l)} \rangle$, l positive integer.

2 $\langle \psi(\alpha x + \beta), f(x) \rangle = |\alpha|^{-1} \langle \psi(x), f(x - \beta/\alpha) \rangle$, $\alpha, \beta \in \mathbf{R}$, $\alpha \neq 0$,

3 $\langle a(x)\psi(x), f(x) \rangle = \langle \psi(x), a(x)f(x) \rangle$.

(ix) Ordinary Functions as Generalised Functions

We now consider an important theorem which is presented without proof and enables us to represent any ordinary function by an equivalent generalised function.

If

1 $f: \mathbf{R} \to \mathbf{C}$,

2 $(1 + x^2)^{-M} |f(x)| \in L(\mathbf{R}, \mathbf{C})$, for some $M \geq 0$

where $L(\mathbf{R}, \mathbf{C})$ is the set of Lesbegue integrable functions, then there is a generalised function $\psi(x) \sim \{f(x;n)\}$ such that

$$\langle \psi, \phi \rangle = \langle f, \phi \rangle, \qquad \phi \in S(\mathbf{R}, \mathbf{C}).$$

In other words, an ordinary function satisfying Condition (2) is equivalent, in the sense of generalised functions, to a generalised function. If, in addition, f is continuous in an interval, then

$$\lim_{n \to \infty} f(x;n) = f(x)$$

is pointwise in that interval. This theorem increases the range of generalised functions available by using not only ordinary functions satisfying Condition (2), but also the new generalised functions which can be obtained by differentiation. A good example of this is the Heaviside step function defined by the expression

$$H(x) = \begin{cases} 1, & \text{if } x > 0; \\ 0, & \text{if } x < 0; \end{cases} \tag{1.8.6}$$

which is a generalised function. In fact, H satisfies the conditions of the theorem with $M = 1$. Moreover, if $\{H(x;n)\}$ is a sequence defining the generalised function H, then for any $\phi \in S(\mathbf{R}, \mathbf{C})$

$$\langle H, \phi \rangle = \int_0^\infty \phi(x)\, dx. \tag{1.8.7}$$

Thus the generalised function H is the process of assigning to a function ϕ a number equal to the area of ϕ from zero to infinity.

In the above example, the condition $\phi \in S(\mathbf{R}, \mathbf{C})$ is a sufficient, but not a necessary, condition to ensure that the integral on the right-hand side exists. Another condition could be

$$\int_0^\infty \phi(x)\, dx < \infty.$$

This is one case in which the set of test functions can be extended to consider test functions which do not belong to $S(\mathbf{R}, \mathbf{C})$.

Now there is sufficient background to consider in detail the most commonly used generalised function, namely the Delta function. The Dirac delta function is defined as

$$\delta(x) \sim \{H'(x;n)\} \tag{1.8.8}$$

where $\{H(x;n)\}$ is the sequence defining the Heaviside step function. δ should be called a delta generalised function instead, but the name delta function is now part of a long tradition. It should be stressed that δ is merely the symbolic representation for all classes of equivalent regular sequences represented by $\{H'(x;n)\}$. Thus

$$\int_{-\infty}^{\infty} \delta(x)\phi(x)\, dx, \qquad \phi \in S(\mathbf{R}, \mathbf{C})$$

actually means

$$\int_{-\infty}^{\infty} \delta(x)\phi(x)\, dx = \langle \delta, \phi \rangle = \lim_{n \to \infty} \int_{-\infty}^{\infty} H'(x;n)\phi(x)\, dx.$$

For the delta function, the set of test functions can be extended to consider bounded and piecewise continuous functions. Since the Heaviside step function defines a generalised function, then for $\phi \in S(\mathbf{R}, \mathbf{C})$ it follows that

$$\begin{aligned}
\langle \delta, \phi \rangle &= \langle H', \phi \rangle \\
&= -\langle H, \phi' \rangle \\
&= -\int_0^\infty \phi'(x)\, dx \\
&= \phi(0).
\end{aligned} \tag{1.8.9}$$

This property of the delta function is called the sampling property and is arguably its most important property which is why some authors define the δ function via its sampling property alone. In other words, we define the delta function in terms of the role it plays in a mathematical operation rather than in terms of what it actually is. Thus, we should always bear in mind that, strictly speaking, the delta function is not really a function even though it is referred to as one. It is actually just one of infinitely many different distributions but its sampling property is unique and is the main reason why it has such a wide range of applications. A more general expression for this property is

$$\begin{aligned}
\langle \delta(x-\alpha), \phi(x) \rangle &= \int_{-\infty}^\infty \delta(x-\alpha)\phi(x)\, dx \\
&= \phi(\alpha).
\end{aligned} \tag{1.8.10}$$

Let $m = 1, 2, 3, \ldots$, then the mth derivative of δ, denoted as $\delta^{(m)}$, is defined symbolically as $\delta^{(m)}(x) \sim \{H^{(m+1)}(x; n)\}$. Thus, there also exists a derivative-sampling property which can be stated as follows:

$$\begin{aligned}
\langle \delta^{(m)}(x-\alpha), \phi(x) \rangle &= \int_{-\infty}^\infty \delta^{(m)}(x-\alpha)\phi(x)\, dx \\
&= (-1)^m \phi^{(m)}(\alpha).
\end{aligned}$$

The main properties of δ are collected together in the following expressions, where $\phi \in S(\mathbf{R}, \mathbf{C})$, $a \in N(\mathbf{R}, \mathbf{C})$, and $\alpha, \beta \in \mathbf{R}$ with $\alpha \neq 0$. These results follow directly from the properties of generalised functions.

$$\begin{aligned}
\langle \delta(-x), \phi(x) \rangle &= \langle \delta(x), \phi(x) \rangle; \\
\langle \delta'(-x), \phi(x) \rangle &= \langle -\delta'(x), \phi(x) \rangle; \\
\langle x\delta(x), \phi(x) \rangle &= \langle 0, \phi(x) \rangle; \\
\langle x\delta'(x), \phi(x) \rangle &= \langle -\delta(x), \phi(x) \rangle; \\
\langle \delta(\alpha x), \phi(x) \rangle &= \langle |\alpha|^{-1}\delta(x), \phi(x) \rangle; \\
\langle \delta(\alpha x + \beta), \phi(x) \rangle &= \langle |\alpha|^{-1}\delta(x + \beta/\alpha), \phi(x) \rangle; \\
\langle a(x)\delta(x), \phi(x) \rangle &= \langle a(0)\delta(x), \phi(x) \rangle.
\end{aligned} \tag{1.8.11}$$

The respective symbolic notation for the above expressions are:

$$\begin{aligned}
\delta(-x) &= \delta(x);\\
\delta'(-x) &= -\delta'(x);\\
x\delta(x) &= 0;\\
x\delta'(x) &= -\delta(x);\\
\delta(\alpha x) &= |\alpha|^{-1}\delta(x);\\
\delta(\alpha x+\beta) &= |\alpha|^{-1}\delta(x+\beta/\alpha);\\
a(x)\delta(x) &= a(0)\delta(x).
\end{aligned} \tag{1.8.12}$$

Note that

$$xf(x) = xg(x) \Longrightarrow f(x) = g(x) + \alpha\delta(x),$$

$$\begin{aligned}
\phi(\alpha) &= \int_{-\infty}^{\infty} \delta(x-\alpha)\phi(x)\,dx\\
&= \int_{-\infty}^{\infty} \delta(\alpha-x)\phi(x)\,dx,
\end{aligned}$$

and

$$\begin{aligned}
\delta(x)\otimes\phi(x) &\equiv \int_{-\infty}^{\infty} \delta(x-\alpha)\phi(\alpha)\,d\alpha\\
&= \phi(x).
\end{aligned}$$

Finally, the following interesting property of the delta function is quoted. Let $f \in C^1(\mathbf{R},\mathbf{C})$ be a function which does not vanish identically and suppose that the roots x_n, $n = 1, 2, \ldots$, of the equation $f(x) = 0$ are such that $f'(x_n) \neq 0$; then

$$\delta[f(x)] = \sum_n \frac{1}{|f'(x_n)|}\delta(x-x_n).$$

Note that, for descriptive purposes only, it is often useful to "visualise" the delta function in terms of the limit of a sequence of ordinary functions or regular sequences $\{\phi(x;n)\}$ which are known as delta sequences, namely

$$\delta(x) = \lim_{n\to\infty} \phi(x;n).$$

1.8.2 Differentiation Across Discontinuities

Let $f \in PC(\mathbf{R},\mathbf{C})$ (a linear space of all $\mathbf{C}$-valued and piecewise continuous functions defined on $\mathbf{R}$) be a function satisfying the conditions of ordinary functions being generalised functions. Let $x_1, x_2, \ldots$, be points at which f is discontinuous and let

$$\sigma_m = f(x_m^+) - f(x_m^-), \qquad m = 1, 2, \ldots,$$

be the jumps of the discontinuities at those points, respectively. The function

$$f_1(x) = f(x) - \sum_m \sigma_m H(x - x_m),$$

where H is the Heaviside step function, is obviously a continuous function everywhere. Moreover, f_1 also satisfies the same conditions as f. Hence f_1 defines a generalised function. Taking derivatives in the sense of generalised functions of both sides of the equation above yields

$$f_1{}'(x) = f'(x) - \sum_m \sigma_m \delta(x - x_m),$$

and therefore it follows that

$$f'(x) = f_1{}'(x) + \sum_m \sigma_m \delta(x - x_m). \tag{1.8.13}$$

Thus, the derivative, in a generalised sense, of a piecewise continuous function across a discontinuity is the derivative of the function in the classical sense plus the summation of the jumps of each discontinuity multiplied by a delta function centred at those discontinuities. This is a generalisation of the concept of a derivative in the classical sense, because if the function were continuous everywhere then σ_m would be zero for all m and the derivative in a generalised sense would coincide with the derivative in the classical sense.

1.8.3 The Fourier Transform of Generalised Functions

The definition of the Fourier transform of a generalised function is based in the following theorem.

If $\phi \in S(\mathbf{R}, \mathbf{C})$ and $\phi(x) \leftrightarrow \Phi(u)$ then $\Phi \in S(\mathbf{R}, \mathbf{C})$ where $\leftrightarrow$ denotes that $\Phi(u)$ is the Fourier transform of $\phi(x)$ which is the inverse Fourier transform of $\Phi(u)$ (so that ϕ and Φ are Fourier transform pairs).

From this result, if $\{\psi(x; n)\} \subset S(\mathbf{R}, \mathbf{C})$ then its Fourier transform $\{\Psi(x; n)\}$ also belongs to $S(\mathbf{R}, \mathbf{C})$. Similarly, if $\phi \in S(\mathbf{R}, \mathbf{C})$, then its Fourier transform $\Phi \in S(\mathbf{R}, \mathbf{C})$ and by Parseval's theorem (see Apostol, 1974)

$$\int_{-\infty}^{\infty} \Psi(u; n)\Phi(u)\, du = \int_{-\infty}^{\infty} \psi(x; n)\phi(x)\, dx$$

so that if the limit

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\psi(x;n)\phi(x)\,dx$$

exists for an arbitrary member of $S(\mathbf{R},\mathbf{C})$, then

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\Psi(u;n)\Phi(u)\,du$$

also exists for any $\Phi\in S(\mathbf{R},\mathbf{C})$. In other words, if $\{\psi(x;n)\}$ is a regular sequence, then so is the sequence $\{\Psi(u;n)\}$.

Let ψ be a generalised function represented by the regular sequence $\psi\sim\{\psi(x;n)\}$, then from the above discussion it follows that the Fourier transform of ψ, denoted by Ψ, is a generalised function represented by the regular sequence $\Psi\sim\{\Psi(u;n)\}$ of the Fourier transform of $\{\psi(x;n)\}$. The inverse Fourier transform can be defined in exactly the same way since $\{\Psi(u;n)\}$ is also a regular sequence.

A result that can also be derived from Parseval's Theorem and from the symmetry of the Fourier transform as follows.

If ψ is a generalised function and ψ and Ψ are Fourier transform pairs and if $\phi\in S(\mathbf{R},\mathbf{C})$ and ϕ and Φ are Fourier transform pairs, then

$$\langle\Psi,\phi\rangle=\langle\psi,\Phi\rangle.$$

From this theorem, using the algebra of generalised functions, it can be shown that the usual properties of the classical Fourier transform are preserved. Also, the following Fourier transform pairs can be derived where H denotes the Heaviside step function; $\alpha,\beta\in\mathbf{R}$ and $m=0,1,2,\ldots$.

$$\begin{aligned}
H(x) &\leftrightarrow \pi\delta(u)+(iu)^{-1};\\
\delta(x) &\leftrightarrow 1;\\
1 &\leftrightarrow 2\pi\delta(u);\\
\delta(x-\alpha) &\leftrightarrow \exp(-iu\alpha);\\
\exp(i\alpha x) &\leftrightarrow 2\pi\delta(u-\alpha);\\
\delta(\alpha x+\beta) &\leftrightarrow |\alpha|^{-1}\exp(-iu\beta/\alpha);\\
\delta^{(m)}(x) &\leftrightarrow (iu)^m;\\
x^m &\leftrightarrow 2\pi i^m\delta^{(m)}(u).
\end{aligned}\tag{1.8.14}$$

The following symbolic integral representation of $\delta(x-\alpha)$ follows directly from the second of the Fourier transform pairs given above;

$$\delta(x-\alpha)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(iux-iu\alpha)\,du.\tag{1.8.15}$$

Finally, the following Fourier transform pair is useful in sampling theory where $0 \neq b \in \mathbf{R}$:

$$\sum_{n=-\infty}^{\infty} \delta(x - nb) \leftrightarrow \frac{2\pi}{b} \sum_{n=-\infty}^{\infty} \delta(u - 2\pi n/b). \tag{1.8.16}$$

The function $\sum\limits_{n=-\infty}^{\infty} \delta(x - nb)$ is called the "comb function" by some authors.

1.8.4 Convolution of Generalised Functions

For generalised functions, the definition of convolution depends on the concept of a direct product; therefore, this section starts by giving the definition of this product.

Let ψ_1 and ψ_2 be generalised functions. The expression

$$\langle \psi_1 \times \psi_2, \phi \rangle \equiv \langle \psi_1(x), \langle \psi_2(y), \phi(x,y) \rangle \rangle, \qquad \phi \in S(\mathbf{R}^2, \mathbf{C}),$$

is called the direct product of ψ_1 and ψ_2.

The direct product is of a particularly simple form when $\phi(x,y)$ is separable. If ψ_1 and ψ_2 are generalised functions and $\phi(x,y) = \phi_1(x)\phi_2(y)$, $\phi_1, \phi_2 \in S(\mathbf{R}, \mathbf{C})$, then

$$\langle \psi_1(x), \langle \psi_2(y), \phi(x,y) \rangle \rangle = \langle \psi_1(x), \phi_1(x) \rangle \langle \psi_2(x), \phi_2(x) \rangle.$$

Thus, for example, the direct product of δ with itself yields the delta function over $\mathbf{R}^2$,

$$\delta(x) \times \delta(y) = \delta(x,y).$$

Over $\mathbf{R}^3$ the result is

$$\delta(x) \times \delta(y) \times \delta(z) = \delta(x,y,z).$$

In order to define the convolution of two generalised functions, let $f, g\colon \mathbf{R} \to \mathbf{C}$ be two functions for which their convolution exists. If ϕ is a function belonging to $S(\mathbf{R}, \mathbf{C})$ then with $\otimes$ denoting the convolution integral,

$$\begin{aligned}
\langle f \otimes g, \phi \rangle &= \int_{-\infty}^{\infty} (f \otimes g)(x)\phi(x)\, dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha) g(x - \alpha)\, d\alpha\, \phi(x)\, dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha) g(y) \phi(\alpha + y)\, d\alpha\, dy
\end{aligned}$$

$$= \int_{-\infty}^{\infty} f(\alpha) \int_{-\infty}^{\infty} g(y)\phi(\alpha + y)\,dy\,d\alpha$$
$$= \langle f(x), \langle g(y), \phi(x+y)\rangle\rangle.$$

In other words, the desired result is equivalent to applying the direct product of f and g to the function $\phi(x+y)$. This result suggests that we define the convolution of two generalised functions ψ_1 and ψ_2 as follows:

$$\begin{aligned}\langle \psi_1 \otimes \psi_2, \phi\rangle &= \langle \psi_1(x), \langle \psi_2(y), \phi(x+y)\rangle\rangle \\ &= \langle \psi_1(x) \times \psi_2(y), \phi(x+y)\rangle, \quad \phi \in S(\mathbf{R}, \mathbf{C}).\end{aligned}$$

Some examples of such convolutions are:

$$\begin{aligned}\delta(x-\alpha) \otimes \psi(x) &= \psi(x-\alpha); \\ \delta^{(m)}(x-\alpha) \otimes \psi(x) &= \psi^{(m)}(x-\alpha); \\ \delta'(x) \otimes H(x) &= \delta(x); \\ \delta(x-\alpha) \otimes \delta(x-\beta) &= \delta(x-\alpha-\beta).\end{aligned}$$

An important consequence of the second of the above results is that every linear differential equation with constant coefficients can be represented as a convolution. Thus if $a_i \in \mathbf{R}$, $i = 0, 1, \ldots, n$, it follows that

$$\sum_{i=0}^{n} a_i f^{(i)}(x) = \left[\sum_{i=0}^{n} a_i \delta^{(i)}(x)\right] \otimes f(x).$$

Note that this statement cannot be made if the convolution operation is restricted to ordinary functions. If ψ_1 and ψ_2 are generalised functions and either ψ_1 or ψ_2 has bounded support then $\psi_1 \otimes \psi_2$ exists.

The convolution of generalised functions preserves the basic properties of the classical convolution except that it is not generally associative. Even if $\psi_1 \otimes (\psi_2 \otimes \psi_3)$ exists as a generalised function it need not to be the same as $(\psi_1 \otimes \psi_2) \otimes \psi_3$; it does not even follow that this will co-exist with $\psi_1 \otimes (\psi_2 \otimes \psi_3)$. An example of this situation is as follows:

$$1 \otimes [\delta'(x) \otimes H(x)] = 1 \otimes \delta(x) = 1,$$

whereas

$$[1 \otimes \delta'(x)] \otimes H(x) = 0 \otimes H(x) = 0.$$

For some classes of generalised functions, e.g. the delta function, it is possible to formulate a convolution theorem. In order to state this theorem the following definitions are needed.

Convergence in $S(\mathbf{R}, \mathbf{C})$

$\{\phi(x;n)\} \in S(\mathbf{R}, \mathbf{C})$ is said to converge in $S(\mathbf{R}, \mathbf{C})$ iff $\{|x|^l \phi^{(m)}(x;n)\}$ converges uniformly over $\mathbf{R}$ for $l > 0$, $m \geq 0$. If the limit function of the sequence $\{\phi(x;n)\}$ is $\phi(x)$ then it may be proved that $\phi \in S(\mathbf{R}, \mathbf{C})$, or the linear space $S(\mathbf{R}, \mathbf{C})$ is closed under convergence.

Multipliers

$\phi: \mathbf{R} \to \mathbf{C}$ is called a multiplier in $S(\mathbf{R}, \mathbf{C})$ iff

1 $\psi \in S(\mathbf{R}, \mathbf{C}) \Longrightarrow \phi \otimes \psi \in S(\mathbf{R}, \mathbf{C})$,

2 $\{\psi(x;n)\} \subset S(\mathbf{R}, \mathbf{C})$, and $\lim_{n\to\infty} \psi(x;n) = 0$
$\Longrightarrow \phi(x)\psi(x;n) \to 0$ as $n \to \infty$ in $S(\mathbf{R}, \mathbf{C})$.

Not all generalised functions are multipliers and so these functions cannot in general be multiplied; the convolution of two arbitrary generalised functions may therefore not necessarily be formed.

The convolution theorem for generalised functions therefore takes the following form:

If

1 ψ_1 and ψ_2 generalised functions,

2 $\psi_1 \leftrightarrow \Psi_1$, $\psi_2 \leftrightarrow \Psi_2$,

3 either ψ_1 or ψ_2 has bounded support,

then either Ψ_1 or Ψ_2 is a multiplier, respectively, and $\psi_1 \otimes \psi_2 \leftrightarrow \Psi_1 \Psi_2$.

The convolution of two arbitrary generalised functions does not always exist. However, in practical problems, such as in electrodynamics, multiplications of arbitrary generalised functions can be used. The formal Fourier transform of these products sometimes give rise to the appearance of convolution integrals which diverge.

1.8.5 The Discrete Representation of the Delta Function

The discrete representation of the delta function $\delta(x - x_0)$ is given by the so-called *Kronecker's delta* δ_{jr} defined by the expression

$$\delta_{jr} = \begin{cases} 1 & \text{if } j = r; \\ 0 & \text{if } j \neq r. \end{cases} \tag{1.8.17}$$

It is a sequence which contains only one non-zero valued element, the value of that element being unity.

The above representation, namely Kronecker's delta, has the discrete-version properties of the delta function. For example, it follows that an arbitrary sequence f_j can be written as a weighted sum of Kronecker delta's:

$$f_j = \sum_{r=-\infty}^{\infty} f_r \delta_{jr}. \tag{1.8.18}$$

This equation is the discrete analogue of the result

$$\phi(\alpha) = \int_{-\infty}^{\infty} \delta(\alpha - x)\phi(x)dx. \tag{1.8.19}$$

Furthermore, δ_{jr} defines the following discrete Fourier transform pairs:

$$\begin{aligned} \delta_{j0} &\leftrightarrow 1; \\ 1 &\leftrightarrow N\delta_{0l}; \\ \delta_{jr} &\leftrightarrow \exp(-2\pi irl/N); \\ \exp(-2\pi ijm/N) &\leftrightarrow N\delta_{jl}. \end{aligned}$$

Finally, the following result is the discrete analogue of the integral representation of the delta function:

$$\frac{1}{N}\sum_{m=0}^{N-1} \exp(2\pi ijm/N)\exp(-2\pi irm/N) = \delta_{jr}.$$

EXERCISES

1.37 By using the sequence $\left\{\frac{n}{\pi}\text{sinc}(nx)\right\}$ where sinc$x$ is the Cardinal sine function $\left(\frac{\sin x}{x}\right)$ show that in the limit as $n \to \infty$

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(ikx)dk.$$

Hence evaluate the Fourier transforms of $\cos x$ and $\sin x$.

1.38 Find the limits (in the distributional sense) of the following

$$\lim_{n\to\infty}\left[\frac{\sin^2(nx)}{nx^2}\right], \qquad \lim_{n\to\infty}[\sin(n!x)].$$

1.39 Show that

$f(x)\delta(x-\alpha) = f(\alpha)\delta(x-\alpha)$; $x\delta(x) = 0$; $\delta(\alpha-x) = \delta(x-\alpha)$; $\delta(\alpha x) = \frac{1}{|\alpha|}\delta(x), \alpha \neq 0$; $\int\limits_{-\infty}^{\infty} f(x)\delta'(x)dx = -f'(0)$; $\int\limits_{-\infty}^{\infty} f(x)\delta^{(n)}(x)dx = (-1)^n f^{(n)}(0)$;

$\delta(\alpha^2 - x^2) = \frac{1}{2\alpha}[\delta(x+\alpha) + \delta(x-\alpha)]$; $\delta(\sin x) = \sum\limits_{n=-\infty}^{\infty} \delta(x - n\pi)$.

1.40 Show that

$$\int\limits_{-\infty}^{\infty} e^{-|x|}\delta(x)dx = e^0 = 1$$

by using the regular sequence $\{n\,\text{gauss}(nx)\}$ and considering

$$\lim_{n\to\infty} \langle n\,\text{gauss}(nx), e^{-|x|}\rangle.$$

1.41 Show that

$\delta(x) \leftrightarrow 1$; $1 \leftrightarrow 2\pi\delta(u)$; $\delta(x-\alpha) \leftrightarrow \exp(-iu\alpha)$; $\exp(i\alpha x) \leftrightarrow 2\pi\delta(u-\alpha)$;

$\delta(\alpha x + \beta) \leftrightarrow \mid \alpha \mid^{-1} \exp(-iu\beta/\alpha)$; $\delta^{(m)}(x) \leftrightarrow (iu)^m$; $x^m \rightarrow 2\pi i^m \delta^{(m)}(x)$.

1.42 The sign function $\text{sgn}(x)$ is defined by

$$\text{sgn}(x) = \begin{cases} 1, & x > 0; \\ -1, & x < 0. \end{cases}$$

By computing the Fourier transform of $\exp(-\epsilon \mid x \mid)\text{sgn}(x)$ over the interval $[-a, a]$ and then letting $a \to \infty$ and $\epsilon \to 0$, show that

$$\text{sgn}(x) \leftrightarrow \frac{2}{iu}$$

given that the Fourier transform of a function $f(x)$ is defined by

$$F(u) = \int\limits_{-\infty}^{\infty} f(x)\exp(-iux)dx.$$

Hence show that

$$H(x) \leftrightarrow \pi\delta(u) - \frac{i}{u}$$

where $H(x)$ is the step function.