Prepared for submission to JHEP

PDEs, ODEs, Analytic Continuation, Special Functions, Sturm-Liouville Problems and All That[1]

C.P. Burgess *Department of Physics, McGill University*

These notes present an introduction to mathematical physics, and in particular the solution of linear ordinary and partial differential equations that commonly arise in physics. Developed in Autumn 1990 for the course Physics 355A at McGill University.

Contents

---

1 Overview

These notes are meant to provide an introduction to several very useful techniques of mathematical physics, which are developed using as a vehicle the very common physical problem of solving boundary-value problems for second-order linear partial differential equations (PDEs). Along the way this requires the development of other very useful tools, including an exploration of the properties of second-order linear ordinary differential equations (ODEs), and in many cases the systematic construction of their solutions.

One of the main lines of development is the construction of series solutions, whose description provides an excuse for a lightning review of several other topics within the calculus of complex variables, including the techniques of analytic continuation that allow the extension to more general domains of solutions initially developed in series form. In particular a description is given of the theory of the kinds of singularities that are possible for solutions of these ODEs, and how to find those solutions that have logarithmic singularities (and so are at first sight not obtainable using series techniques).

The extension of solutions in this way, and the special functions to which this leads, are presented in a more unified way than is often done. Rather than regarding each type of special function on its own terms, with properties to be developed on a case-by-case basis, the notes instead classify the kinds of differential equations that typically arise in physical applications. In particular, since these usually lead to equations involving three (or fewer) regular-singular points (RSPs, whose precise definition is given when appropriate below), the most general form of this equation is identified and solved once and for all. Since the most general ODE of this type is the Hypergeometric equation, the properties of the solutions to this equation — Hypergeometric functions — are studied in some detail. It is because so many physical systems involve equations with three or fewer RSPs that most of the special functions usually studied are special cases of these Hypergeometric functions, and this is why their properties follow as special cases of those of the Hypergeometric functions.

Also discussed in detail are the Confluent Hypergeometric equations and functions, that are obtained when two of the RSPs happen to coalesce to give an ODE with one regular and one irregular singular point. The properties of these functions (and their many special cases of physical interest, such as Bessel functions) are also obtained as limiting instances of the general Hypergeometric case.

The main bridge from solutions to ODEs to solutions of PDEs comes through the technique of separation of variables, which constructs solutions in product form — such as $\psi(x,y) = X(x)Y(y)$. Of course most solutions to PDEs do *not* have this form, so the key idea is that a general solution to the PDE can be written as a linear combination of solutions in this product form. This leads to the discussion of Sturm-Liouville problems and the use of special functions as bases for the expansion of more general functions — including Fourier series, Fourier-Bessel series and so on. This brings out the relationship between calculus and linear algebra, in a way that connects the treatments of classical electrodynamics and quantum mechanics that undergraduates usually study at the same time as mathematical physics.

These are old subjects that at one time were part of the standard lore for physicists, but some of them are less often taught nowadays due to the advent of cheap numerical methods. Although such numerical methods are of course very useful (and should always be exploited), it behooves the well-trained physicist also to know these more analytic methods. They are useful both to understand how the numerical methods work, but also to allow one to check one's numerics by providing analytic comparisons in various limits.

A Road Map

The presentation of the remainder of the notes proceeds as follows. The first few sections describe the types of linear second-order PDEs that often arise in physics, and argue why it is that the types of equations encountered are often so similar in physical applications. The procedure of separation of variables is then described in a few simple examples.

There follows then a lengthy discussion about constructing solutions to linear second-order ODEs. This involves first describing general properties of the vector space of their solutions, and then moves on to the discussion of the construction of series solutions as well as of the limitations of this procedure. Along the way is a telegraphic summary of some useful results in the calculus of a single complex variable, including contour integration and analytic continuation, since these are often missing in the background of an undergraduate learning this material.

Next comes the classification of the ODEs arising in physical applications, and the demonstration that most problems encountered fall into the class described by Hypergeometric functions. The properties of these are then enumerated, followed by a discussion of how the usual special functions (Bessel, Legendre, Gegenbauer, Hermite, Laguerre, etc) arise as special cases of the Hypergeometric functions.

The penultimate sections explore how the space of solutions to a linear PDE is a vector space and how separated solutions involving the above the special functions can be used to find a basis of this vector space, the bailiwick of Sturm-Liouville problems. All these techniques are tied together in the final section which uses them to solve several explicit example PDE examples from start to finish.

1.1 Introduction

Consider a physical variable, $\psi(x, y, z, t)$, that varies throughout space, (x, y, z), and time, t. ψ could represent the pressure, density or temperature of a fluid, or the value of the electromagnetic scalar potential, or any number of other continuous physical variables. The evolution of such an object in time is generally described by a *partial differential equation*, or P.D.E.. These equations are the analogue for a continuously varying variable of the usual ordinary differential equations describing the motion of a particle. For instance Newton's laws for the motion of an element of fluid might take the form:

$$\frac{\partial^2 \psi}{\partial t^2} + F\left[\psi, \frac{\partial \psi}{\partial t}, \nabla\psi, ...\right] = 0. \tag{1.1.1}$$

in which F gives the 'force' acting on ψ at every point as a function of the behaviour of ψ in the neighbourhood of that point. Different physical situations will be described by different functional forms for F.

From a mathematical point of view, a physicist's role in life is to find which P.D.E.'s describe which physical problems (*i.e.* which functions, F) and then to construct and interpret the corresponding solutions. For this reason a great deal of effort in theoretical physics and applied mathematics is and has been devoted to the solving of these differential equations.

F can be some horribly complicated function of its arguments, and in this general case only very little may be said about the properties of the solutions to eq. (1.1.1). A problem that is much better understood, and which nevertheless has applications to many physical situations, is the case where ψ and its derivatives are in some sense 'small'. In this case one can imagine expanding F in powers of its arguments as follows:

$$F\left[\psi, \frac{\partial \psi}{\partial t}, \nabla\psi, ...\right] \approx -f + C\psi + \sum_{i=1}^{4} B_i \frac{\partial \psi}{\partial x_i} + \sum_{i=1}^{4}\sum_{j=1}^{4} A_{ij} \frac{\partial^2 \psi}{\partial x_i \partial x_j} + \cdots . \tag{1.1.2}$$

The coefficient functions f, C, B_i and A_{ij} may all depend on the coordinates x, y, z and t. The ellipsis (*i.e.* '$\cdots$') denote terms involving higher derivatives and/or higher powers of ψ and its derivatives. These may be neglected when ψ and its derivatives are sufficiently small.

Neglecting these terms and using eq. (1.1.2) in the P.D.E. (1.1.1) gives a differential equation that is *linear* in ψ and at most *second order* in its derivatives. A good deal more is known about the construction of solutions to these problems.

These lectures are devoted to solving the linear, second-order P.D.E.'s that arise most frequently in mathematical physics. Having said this, less than half of the course (chapters 4, 5 and 12) is actually spent directly solving these P.D.E.'s. The remaining chapters (6 through 11) are devoted to a long digression on generating solutions to the linear second-order *ordinary* differential equations (O.D.E.'s) and to describing their solutions in some detail.

A major reason for this digression is that much of the theory of P.D.E.'s can be constructed by using the corresponding theory of O.D.E.'s as a guide. This line of argument is fleshed out in Chapters 3 and 4.

Furthermore, the principal method of solution discussed here for P.D.E.'s is the method of *Separation of Variables* (Chapter 5) in which the partial differential problem is reduced to the construction of solutions to a related set of O.D.E.'s. This technique consists of the construction of solutions with a specific dependence upon the independent variables x, y, z and t:

$$\psi(x, y, z, t) = X(x)Y(y)Z(z)T(t). \tag{1.1.3}$$

This would superficially appear to have limited utility since the solution to a generic physical problem does not have this *separated* form. The key observation is then that although the general solution does not have this separated form it is often possible to find a *basis* of solutions with this form. Any solution may then be written as a linear combination of this basis of solutions. The construction of this basis is performed in Chapters 8 through 10 and the claim that any solution may be expanded in terms of the basis is the topic of Chapter 11. Chapter 12 is then a summary in which all of the parts of the argument are pulled together and applied to specific boundary-value problems.

A completely different reason for the lengthy detour through the theory of O.D.E.'s is that many of these topics, such as *Analytic Continuation* and *Steepest Descent* (Chapters 6 and 7), *Special Functions* (Chapters 9 and 10), and *Self-Adjoint Eigenvalue Problems* (Chapter 11), are useful in their own right elsewhere in mathematical physics. The problem of solving the P.D.E.'s is being partially used here as the vehicle through which they are presented.

2 Partial Differential Equations of Mathematical Physics

Our goal is to find solutions to the 2nd order, linear P.D.E.'s that arise in mathematical physics. This section first describes the P.D.E.'s that commonly arise in physical problems and then gives an illustrative derivation of one such P.D.E.. Some arguments are presented as to why the P.D.E.'s of interest take the form that they do.

2.1 Sample P.D.E.'s

The most general 2nd order, linear, inhomogeneous P.D.E. governing a variable, ψ, that depends on the coordinates x_k, $k = 1, .., 4$ with $x_1 = x$, $x_2 = y$, $x_3 = z$, and $x_4 = t$, is:

$$\mathcal{L}\psi = f.$$

Here $f = f(x_k)$ is a known function of the coordinates, x_k, and $\mathcal{L}$ is the following differential operator:

$$\mathcal{L}\psi = \sum_{i=1}^{4}\sum_{j=1}^{4} A_{ij}(x_k)\frac{\partial^2\psi}{\partial x_i \partial x_j} + \sum_{i=1}^{4} B_i(x_k)\frac{\partial\psi}{\partial x_i} + C(x_k)\psi \tag{2.1.1}$$

The coefficients, A_{ij}, B_i, and C are, like f, all given functions of the coordinates only and *not* of ψ itself. It is conventional to choose A_{ij} to be symmetric $A_{ij} = A_{ji}$, as may always be done without loss of generality.

Solutions of the general P.D.E. are not known so it is fortunate that the most general form given by eq.(2.1.1) does not often (if ever) arise in mathematical physics. Those P.D.E.'s that do arise most frequently are listed below:

1. Laplace's Equation:

   Many time-independent problems, such as the description of an electrostatic potential in the absence of charges, are described by *Laplace's equation*. This is defined for $\psi = \psi(x, y, z)$ by:

   $$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = 0. \tag{2.1.2}$$

   The differential operator, ∇^2, defined by eq.(2.1.2) is called the *Laplacian* operator, or just the *Laplacian* for short.

2. Poisson's Equation:

   The inhomogeneous version of Laplace's equation is called *Poisson's equation*. It governs the behaviour of an electrostatic potential in the presence of charge distributions, for example. It has the form:

   $$\nabla^2\psi = f(x, y, z) \tag{2.1.3}$$

   in which f is a known function.

3. Helmholtz's Equation:

   The equation:
   $$\nabla^2\psi + k^2\psi = 0 \tag{2.1.4}$$
   is known as *Helmholtz's equation*. k is a constant. This equation arises in, among other places, the study of the propagation of waves.

4. Schrödinger's Equation:

   A great deal of quantum mechanics is devoted to the study of the solutions to the *time-dependent Schrödinger equation*:
   $$-\frac{\hbar^2}{2m}\nabla^2\psi + V(x,y,z)\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{2.1.5}$$
   This equation governs the time dependence of the wave-function of a particle moving in a given potential, $V(x,y,z)$. A special role is played by solutions to (2.1.5) that have the simple form: $\psi = \phi(x,y,z)\exp(-iEt/\hbar)$. The function ϕ satisfies the *time-independent Schrödinger equation*:
   $$-\frac{\hbar^2}{2m}\nabla^2\phi + V(x,y,z)\phi = E\phi. \tag{2.1.6}$$
   In both of these equations $\hbar$, m, and E represent real constants. i, as usual, satisfies $i^2 = -1$.

5. The Diffusion Equation:

   The P.D.E. governing the diffusion of a quantity, such as the number of particles per unit volume present in a region, is the *diffusion equation*:
   $$\frac{\partial\psi}{\partial t} - \kappa\nabla^2\psi = 0. \tag{2.1.7}$$
   $\kappa = k^2$ denotes here a real constant.

6. The Wave Equation:

   Wave propagation, including pressure waves in fluids, electromagnetic waves and gravitational waves all satisfy the following *wave equation*:
   $$-\frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} + \nabla^2\psi = 0. \tag{2.1.8}$$
   The real constant v can be interpreted as the speed of the corresponding wave.

7. The Klein-Gordon Equation:

   Disturbances travelling through fields that mediate forces with a finite range, satisfy a modification of the wave equation called the *Klein-Gordon equation*. It is given by:
   $$-\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} + \nabla^2\psi + \frac{m^2c^2}{\hbar^2}\psi = 0. \tag{2.1.9}$$
   The coefficients c, $\hbar$ and m all represent constants.

There are several features that these P.D.E.'s have in common:

1. The techniques used to solve these equations rely on a property that they all share. This common property is the feature that they are all linear in the dependent variable, ψ. Although the theory of linear equations has received the bulk of attention from applied mathematicians and theoretical physicists, it is by no means true that most of the P.D.E.'s that arise in real physical problems are linear. The prominence of linear equations in this list reflects the fact that these equations can be generally solved whereas comparable methods of solution for nonlinear equations are relatively scarce.

   In some physical situations, however, linear equations provide an adequate approximation to a nonlinear problem. Nonlinear terms in a differential equation are those that involve quadratic and higher powers of the field, ψ, or its derivatives. If ψ and its derivatives are small enough then its square and higher powers may be expected to be smaller still. In this case any linear terms in a P.D.E. may be expected to dominate the nonlinear ones. Besides the advantage of being well understood, linear P.D.E.'s are therefore also of interest as an approximation to the full nonlinear behaviour in the limit of weak, slowly varying fields.

2. As pointed out in the introduction, these P.D.E.'s all involve at most two derivatives with respect to the independent variables, x, y, z and t. This usually arises from the physical requirement of stability since equations involving higher time derivatives generically admit runaway solutions. In some cases the physics of interest does lead to equations involving higher than second derivatives. In these cases the neglect of the higher-derivative terms is justified by the approximation that the fields in the problem are slowly varying in space and time.

3. All of the above equations have two more features in common. Firstly, there are absolutely no terms that involve just a single spatial derivative. Secondly, the second derivative of ψ with respect to the spatial coordinates always appears in the combination $\nabla^2\psi$. This is a reflection of the underlying symmetry of the corresponding problems with respect to spatial rotations. A similar observation holds for the combination of derivatives that appears in the wave equation and in the Klein-Gordon equation. These can be derived by requiring that they be invariant with respect to the Lorentz transformations of special relativity.

4. With the exception of the two Schrödinger equations all of the homogeneous P.D.E.'s listed involve only constant coefficients. This follows from the invariance of the underlying physical situations under translations in space and time.

2.2 A Typical Derivation

It is useful to have in mind some idea of how these equations arise in physical situations because physical intuition about the underlying problem is a useful guide when trying to derive the properties of their solutions. We therefore now turn to a simple derivation of one of the above equations. We take the diffusion equation as our example.

Consider a system contained within some volume, R, consisting of an ideal gas that is *locally* in thermal equilibrium. Local equilibrium means that each small volume element of the medium is assumed to be in equilibrium with its immediate surroundings, allowing thermodynamic variables such as temperature, $T(x, y, z, t)$, and energy per unit volume, $U(x, y, z, t)$, to be defined at each point. For an ideal gas these are related by: $U = c_v T$ in which c_v is the *specific heat* at constant volume and is independent of x, y, z and t. Thermodynamic stability requires that c_v be positive. We wish to understand how an initially inhomogeneous temperature distribution will diffuse throughout R.

Consider an arbitrary volume, V, contained within R. The first law of thermodynamics relates the rate of change of the energy contained within V to the *rate* of heat flow (*i.e.* heat flux) into V. Denoting the energy within V as

$$E = \int_V U \, dV = \int_V c_v T \, dV, \tag{2.2.1}$$

the first law states:

$$\frac{dE}{dt} = \frac{\delta Q}{\delta t}. \tag{2.2.2}$$

$\delta Q/\delta t$ denotes the total heat flux into V. In this system heat moves about as a result of the motion of the gas molecules. The heat flux per unit time through any surface element, $\mathbf{n} \, dS$, can be described in terms of a heat-flux vector: $\mathbf{q}(x, y, z, t)$, defined throughout R. Here $\mathbf{n}(x, y, z)$ denotes the unit normal to the surface element and dS is its infinitesimal area. The flux through the surface element is given in terms of $\mathbf{q}$ by $\mathbf{q} \cdot \mathbf{n} \, dS$.

The heat flux into V is then given by:

$$\frac{\delta Q}{\delta t} = -\int_{\partial V} \mathbf{q} \cdot \mathbf{n} \, dS \tag{2.2.3}$$

in which ∂V denotes the boundary of V and $\mathbf{n}$ is the outward-pointing unit normal to the boundary. Combining eqs.(2.2.2) and (2.2.3) allows the first law to be written:

$$c_v \int_V \frac{\partial T}{\partial t} \, dV = -\int_{\partial V} \mathbf{q} \cdot \mathbf{n} \, dS. \tag{2.2.4}$$

Using the divergence theorem:

$$\int_{\partial V} \mathbf{q} \cdot \mathbf{n} \, dS = \int_V \nabla \cdot \mathbf{q} \, dV \tag{2.2.5}$$

then gives:

$$\int_V \left(c_v \frac{\partial T}{\partial t} + \nabla \cdot \mathbf{q} \right) dV = 0, \qquad \text{for all } V \subset R. \tag{2.2.6}$$

Since the integral over V in eq.(2.2.6) must vanish for any choice of V that lies within R, it must be true that the integrand itself vanishes throughout R if it is sufficiently smooth. To see this argue as follows: We know that $\int_V f \; dV = 0$ for all $V \subset R$ and we wish to argue that $f(p) = 0$ for any $p \in R$. The argument proceeds by contradiction. Assume therefore that $f(p)$ is not zero for some $p \in R$. $f(p)$ must then be either positive or negative. If $f(p) > 0$ then, if f is continuous, f must also be positive throughout some neighbourhood, N, containing p. Choose, then, $V \subset N \subset R$ so that f is positive throughout the entire region of integration. Since the integral of a positive function is itself positive it must be that, for this V, $\int_V f \; dV > 0$ in contradiction with the assumption that this integral vanishes. The argument is identical if $f(p)$ is instead assumed to be negative. This lemma therefore allows us to conclude from eq.(2.2.6) that:

$$c_v \frac{\partial T}{\partial t} + \nabla \cdot \mathbf{q} = 0 \tag{2.2.7}$$

throughout R. This is the local expression of the physical content of the first law: conservation of energy.

In order to proceed further we need another piece of information that relates T to $\mathbf{q}$. This is provided by a phenomenological relation of the form:

$$\mathbf{q} = \mathbf{q}[T, \nabla T, ...]. \tag{2.2.8}$$

This expresses how the transport of energy is governed by the properties of the surrounding gas. It could in principle be derived from the equations of motion of the underlying gas molecules. If, however, T is assumed to be varying slowly throughout space we expect that $\mathbf{q}$ should be well approximated by those terms that involve the fewest derivatives of T. Since $\mathbf{q}$ is a vector quantity so must be the right-hand-side of eq.(2.2.8). There is only one term possible that both transforms like a vector and involves the fewest number of derivatives of T and this is:

$$\mathbf{q} \approx -\lambda^2 \nabla T. \tag{2.2.9}$$

The sign of the right-hand-side of eq.(2.2.9) is chosen so that $\mathbf{q}$ points in the direction of *decreasing* temperature, as would be expected physically. The real constant λ can be calculated in principle from an understanding of the motion of the underlying molecules.

Equations (2.2.7) and (2.2.9) may now be used to elimate $\mathbf{q}$ and so derive an equation involving just T. Taking the divergence of eq.(2.2.9), and using the identity $\nabla \cdot \nabla T = \nabla^2 T$, gives the diffusion equation:

$$\frac{\partial T}{\partial t} - \kappa \nabla^2 T = 0. \tag{2.2.10}$$

The *diffusion constant*, κ, is found to be related to λ and c_v by $\kappa = \lambda^2 / c_v$. This is a manifestly positive quantity.

The diffusion equation has many solutions. Being a linear homogeneous equation indeed implies that the sum of any two solutions is itself a solution. The real temperature distribution occuring in any physical problem, however, satisfies both the diffusion equation and a set of boundary conditions. The above derivation gives some physical intuition for what kinds of conditions would be necessary to specify the solution uniquely—that is to say, what constitutes a *well-posed* boundary-value problem for the diffusion equation.

Physically, the behaviour of the temperature in R should depend on both the initial temperature distribution and the amount of heat flux, if any, that passes through the boundary, ∂R, of R. We expect, therefore, that a well-posed boundary-value problem for the diffusion equation would be to find the temperature distribution, $T(x, y, z, t)$, satisfying:

$$
\begin{aligned}
\frac{\partial T}{\partial t} &= \kappa \nabla^2 T \qquad \text{for all } (x, y, z) \in R \qquad \text{and for all } t \\
\text{and:} \qquad T &= \tau(x, y, z) \qquad \text{for all } (x, y, z) \in R \qquad \text{and for } t = 0 \\
\text{and:} \qquad \mathbf{n} \cdot \nabla T &= f(x, y, z, t) \qquad \text{for all } (x, y, z) \in \partial R \qquad \text{and for all } t
\end{aligned}
$$

We prove in the following sections that these conditions do indeed guarantee a unique solution to this equation.

3 General Properties of Ordinary Differential Equations

3.1 Introduction

Our goal is to construct solutions to P.D.E.'s such as those listed in the previous section. In order to do so we must answer the following general questions:

1. What is the general solution?
2. What constitutes a well-posed boundary-value problem?
3. How is the general solution constructed?
4. Given the general solution, what is the particular solution satisfying a given set of boundary conditions?

Our strategy in answering these questions is to use experience with ordinary differential equations, O.D.E.'s, as a guide. The method used to construct the solutions of the P.D.E.'s also consists of reducing the problem to one of solving an equivalent set of O.D.E.'s. This chapter is therefore devoted to a review of the answer to the analoguous problems in the theory of linear 2nd-order O.D.E.'s. The explicit construction of, and a discussion of the resulting properties of, their solutions is deferred to later chapters.

3.2 The Space of Solutions

The general form for a linear, 2nd-order O.D.E. is:

$$a(x)y'' + b(x)y' + c(x)y = g(x). \tag{3.2.1}$$

where a, b, c and g are given real functions of the real independent variable, x, and a vanishes only at isolated points in the interval in which x takes its values. (If a were to vanish on some interval the problem to be solved in this interval would be a *first-order* O.D.E. rather than a second-order one.) It is conventional to divide the equation through by $a(x)$ giving:

$$\mathcal{L}y = f(x)$$

$$\text{with:} \qquad \mathcal{L}y = y'' + p(x)y' + q(x)y. \tag{3.2.2}$$

The new coefficient functions are related to the previous ones in an obvious way.

The fundamental property that underlies the approach to solving this O.D.E. is the *linearity* of the differential operator $\mathcal{L}$. By definition linearity is the statement that if $y_1(x)$ and $y_2(x)$ are any two functions and α and β are real (or complex) numbers, then:

$$\mathcal{L}(\alpha y_1 + \beta y_2) = \alpha \mathcal{L}y_1 + \beta \mathcal{L}y_2. \tag{3.2.3}$$

The utility of this property follows from its wide-ranging consequences for the space of solutions to these O.D.E.'s:

1. If y_1 and y_2 are both solutions to the O.D.E. $\mathcal{L}y = f$, with $\mathcal{L}$ given in eq.(3.2.2), then $y_h = y_2 - y_1$ satisfies the *homogeneous* equation:

$$\mathcal{L}\, y_h = 0. \tag{3.2.4}$$

   This implies that an arbitrary solution, y_2, to (3.2.2) is given by any particular solution, y_1, plus a solution to the homogeneous equation, (3.2.4). The problem of finding the general solution to the inhomogeneous equation is therefore reduced to the much simpler problem of finding a single particular solution together with the problem of finding the general solution to the homogeneous equation. We return to the construction of particular solutions to the inhomogeneous equation in chapter 13. Properties of solutions to the homogeneous problem are the subject of the rest of this section.

2. Linearity also ensures that the space of solutions, $\mathcal{S}$, to the homogeneous O.D.E., (3.2.4), defined on some real interval $[a, b]$, forms a vector space. That is, $\mathcal{S}$ together with the two operations defined by:

$$\begin{array}{rc} \text{pointwise addition:} & [y_1 + y_2](x) \equiv y_1(x) + y_2(x) \\ \text{scalar multiplication:} & [\alpha y](x) \equiv \alpha[y(x)] \end{array}$$

   forms a vector space. The 'zero vector' of this vector space is the function that vanishes for all x in $[a, b]$.

   The proof that these definitions of 'vector addition' and 'scalar multiplication' give $\mathcal{S}$ a vector-space structure consists of a straightforward check of the defining properties of a vector space. (See, for example, *Topics in Algebra* (Wiley).) Satisfaction of these properties relies crucially, of course, on the linearity of $\mathcal{L}$.

   The significance of this fact is that it should therefore be possible to write any of the elements of $\mathcal{S}$, y say, as a linear combination of a *basis* of solutions. $y(x) = \sum_k c_k y_k(x)$ with the c_k's being constants. The general solution to $\mathcal{L}y = 0$ is then known once a basis of solutions is constructed.

In order to sharpen this logic recall the following definitions.

Definition 3.1. *A set of vectors, $\{y_1, ..., y_n\}$, is linearly independent if and only if the equation $c_1 y_1 + ... + c_n y_n = 0$ implies that all of the constants, $\{c_1, ..., c_n\}$, must vanish. The set, $\{y_1, ..., y_n\}$, is said to be linearly dependent if it is not linearly independent. This is equivalent to the existence of a set of constants, $\{c_1, ..., c_n\}$, not all vanishing, for which $c_1 y_1 + ... + c_n y_n = 0$.*

Definition 3.2. *A vector space has dimension n if and only if both of the following are true: (i) a set of n linearly independent vectors exists, and (ii) any set of $n + 1$ vectors is linearly dependent.*

Definition 3.3. *A basis of an n-dimensional vector space is any set of n linearly-independent vectors.*

Clearly these definitions imply that if $\{y_1, ..., y_n\}$ forms a basis, then any other vector, y, can be written as the sum $y = \sum_k c_k y_k$. This follows from the fact that the $n+1$ vectors, $\{y_1, ..., y_n, y\}$, must be linearly dependent. In the event that the vector space should prove to be infinite dimensional a basis can be defined by this last condition that an arbitrary vector can be expressed as a linear combination of basis vectors.

In practice, for the space of solutions, $\mathcal{S}$, of $\mathcal{L}y = 0$ we need to answer the following questions:

1. What is the dimension of $\mathcal{S}$?
2. What constitutes a basis for $\mathcal{S}$?
3. What kind of boundary conditions uniquely specify an element of $\mathcal{S}$?

The remainder of this section is devoted to (partially) answering the problems in this list. We first show that if any nonzero solution exists, then $\mathcal{S}$ is two-dimensional. We then demonstrate how to construct from any nonzero solution a second linearly independent solution. These two elements of $\mathcal{S}$ form a basis. We finally show that the initial-value problem is well posed, and so uniquely specifies a solution to the O.D.E.. The existence and explicit construction of the first nonzero solution that these results assume is addressed in chapter 8.

3.3 The Wronskian

Before proceeding to the three results listed in the previous section a detour is necessary to derive a more useful criterion for the linear independence of n functions. Recall that the definition of linear independence, when applied to the solution space, $\mathcal{S}$, is the following statement:

$$c_1 y_1(x) + ... + c_n y_n(x) = 0 \quad \text{for all } x \implies c_1 = c_2 = ... = c_n = 0. \tag{3.3.1}$$

Since the left-hand-side of this implication is true for all x it may be differentiated. Successive differentiation shows that (3.3.1) is *equivalent* to:

$$\left.\begin{array}{c} c_1 y_1(x) + ... + c_n y_n(x) = 0 \\ c_1 y_1'(x) + ... + c_n y_n'(x) = 0 \\ \vdots \\ c_1 y_1^{(n-1)}(x) + ... + c_n y_n^{(n-1)}(x) = 0 \end{array}\right\} \quad \text{for all } x \implies c_1 = c_2 = ... = c_n = 0. \tag{3.3.2}$$

The left-hand-side has the form of a set of n linear equations in the n unknowns $\{c_1, ..., c_n\}$. Such a system of equations has $c_1 = c_2 = ... = c_n = 0$ as its only solution if and only if the following determinant does not vanish:

$$W_n(x) \equiv \begin{vmatrix} y_1(x) & \dots & y_n(x) \\ y_1'(x) & \dots & y_n'(x) \\ \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & \dots & y_n^{(n-1)}(x) \end{vmatrix}. \tag{3.3.3}$$

This last equation defines the *Wronskian* of the n functions $\{y_1(x), ..., y_n(x)\}$. In terms of $W_n(x)$, the functions $\{y_1(x), ..., y_n(x)\}$ are linearly independent if and only if there exists a point, $x_0 \in [a, b]$, for which $W_n(x_0) \neq 0$. Conversely, the $\{y_k(x)\}$ are linearly dependent if and only if $W_n(x) = 0$ for all $x \in [a, b]$.

With this test for linear dependence in hand we now turn to the discussion of $\mathcal{S}$.

Theorem 3.1. *Any three solutions, $y_1(x)$, $y_2(x)$ and $y_3(x)$, of the differential equation $\mathcal{L}y = 0$, with $\mathcal{L}$ given by eq.(3.2.2) are linearly dependent.*

Proof. To prove this theorem consider the Wronskian:

$$W_3(x) = \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix}. \tag{3.3.4}$$

The differential equation (3.2.4) implies that $y_k'' + p(x)y_k' + q(x)y_k = 0$ for $k = 1, 2$ and 3. This implies that the bottom row of the matrix in eq.(3.3.4) is a linear combination of the first two rows. Its determinant, $W_3(x)$, must therefore vanish for all x, as may be checked by explicit evaluation. The three functions are therefore linearly dependent as claimed. □

This theorem tells us that the space of solutions to the O.D.E. (3.2.4) is at most two-dimensional. We now show that if $\mathcal{S}$ has any nonzero element at all, then it is exactly two-dimensional.

Theorem 3.2. *If $y_1(x) \not\equiv 0$ is a solution to the O.D.E. (3.2.4): $\mathcal{L}y_1 = 0$, then a linearly-independent solution, $y_2(x)$ exists. (The construction of the original solution, $y_1(x)$, is deferred to a later section.)*

Proof. This theorem is proven by explicitly constructing the second solution. Consider the Wronskian of any two solutions to the O.D.E. (3.2.4):

$$W_2(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'. \tag{3.3.5}$$

Eq.(3.2.4) implies that $W_2(x)$ satisfies its own differential equation:

$$\frac{dW_2}{dx} = y_1 y_2'' - y_2 y_1'' = -p(x)[y_1 y_2' - y_2 y_1'] = -p(x) W_2(x). \tag{3.3.6}$$

This is easily integrated to yield the solution:

$$W_2(x) = W_2(x_0) \exp\left(-\int_{x_0}^{x} p(u)\, du\right). \tag{3.3.7}$$

This last equation implies that once W_2 is known at a single point, its value is determined at all other points solely by the O.D.E.. The idea is to now regard eq.(3.3.5) as being an equation from which y_2 is to be determined in terms of a known function y_1 and W_2 given by eq.(3.3.7). To solve rewrite eq.(3.3.5):

$$y_1^2 \frac{d}{dx}\left(\frac{y_2}{y_1}\right) = W_2(x) = A \exp\left(-\int_{x_0}^{x} p(u)\, du\right),$$

divide by y_1^2, and integrate:

$$y_2(x) = y_1(x)\left[B + A\int^x du\, \frac{1}{y_1^2(u)} \exp\left(-\int^u dv\, p(v)\right)\right]. \tag{3.3.8}$$

It is easily checked that $y_2(x)$ as defined by eq.(3.3.8) satisfies the O.D.E. (3.2.4) and has a Wronskian with $y_1(x)$ given by $W_2 = A\exp\left(-\int^x p(u)\, du\right)$. Given any solution to eq.(3.2.4), then, eq.(3.3.8) defines a second linearly independent solution provided only that the integrals in eq.(3.3.8) converge. This is the result that was to be proven. □

Given the existence of any nontrivial solution to the O.D.E. the last two theorems imply that the space, $\mathcal{S}$, of solutions is two-dimensional. Any nonzero solution, y_1, together with the second solution, y_2, defined by eq. (3.3.8) gives a basis for $\mathcal{S}$. This in turn implies that the general solution to (3.2.4), $y_h(x)$ can be written in terms of these two basis solutions by $y_h(x) = c_1 y_1(x) + c_2 y_2(x)$. The problem of finding the general solution to the O.D.E. (3.2.4) has, in principle, boiled down to the construction of a single nonzero solution. As discussed earlier, construction of such a solution also gives the general solution to the inhomogeneous equation, eq. (3.2.2), when combined with any particular integral of that equation.

3.4 Initial-Value Problems

Since the general solutions of (3.2.2) and (3.2.4) involve two free constants, c_1 and c_2, we expect to require two pieces of boundary information in order to get a unique solution. It is important to realize, however, that just any two pieces of information need not make the solution unique. Consider, for example, the following O.D.E. and boundary-value problem:

$$\frac{d^2y}{dx^2} + y = 0 \qquad \text{for } \ x \in [0, \pi]$$

$$\text{with:} \qquad y(0) = y(\pi) = 0.$$

This problem has the one-parameter family of solutions: $y(x) = a \sin x$.

There is a general lesson here. The reason this choice of boundary conditions proved insufficient to determine the solution was that it, like the O.D.E. itself, was linear and homogeneous. That is, any linear combination of solutions to the O.D.E. *and* the boundary conditions is also a solution of both. This implies that if any nonzero solution, $y(x)$, exists then so does the one-parameter family $\alpha y(x)$ for any real constant α.

Theorem 3.3. *One choice of boundary conditions that is always guaranteed to give a unique result for the O.D.E. (3.2.4) is the initial-value problem:*

$$y(x_0) = u \qquad \text{and} \qquad \frac{dy}{dx}(x_0) = v \tag{3.4.1}$$

with u and v not both zero.

Proof. To see that the solution to (3.2.2) with initial condition (3.4.1) is unique suppose that two solutions, y_1 and y_2 exist. Their Wronskian is given by eq.(3.3.7) with $W_2(x_0) = y_1(x_0)y_2'(x_0) - y_2(x_0)y_1'(x_0)$ which vanishes by virtue of eq.(3.4.1). This implies that there are two constants, c_1 and c_2 not both zero, for which $c_1 y_1(x) = c_2 y_2(x)$ for all $x \in [a, b]$. Evaluating this equation at $x = x_0$ then implies $c_1 = c_2 \neq 0$ unless u vanishes. If $u = 0$ then v cannot vanish so the argument may be repeated using $c_1 y_1'(x) = c_2 y_2'(x)$.

The case $u = v = 0$ is trickier since in this case the O.D.E. and boundary conditions are both linear and homogeneous. Whether or not the initial-value problem is well-posed depends in this case on how smooth the solutions are required to be. If the solutions are required to be analytic on $[a, b]$ then the solution is unique (and is identically zero). To see this notice that $y(x_0) = y'(x_0) = 0$ together with the O.D.E. (3.2.4) implies that $y''(x_0) = 0$. Similarly, repeated differentiation of the O.D.E. shows that all derivatives of y vanish when evaluated at x_0. This implies that every term in the Taylor expansion of y about $x = x_0$ vanishes (see section 6.2), so $y = 0$ throughout the interval $[a, b]$. The zero function is therefore the unique analytic solution to the initial-value problem.

If, on the other hand, the solution to (3.2.4) with $y(x_0) = y'(x_0) = 0$ need not be analytic then it need not be unique. To see this consider the following sample O.D.E.:

$$\frac{d^2y}{dx^2} - \frac{1}{x^2}\frac{dy}{dx} + \frac{2}{x^3}y = 0 \qquad \text{for } x \in [0, \infty).$$

A one-parameter family of solutions to this equation and the initial conditions $y(0) = y'(0) = 0$ is given by $y = a \exp(-1/x)$. Although $y(x)$ is here infinitely differentiable it is not analytic at $x = 0$. □

Putting aside the properties of O.D.E.'s, we turn now to a consideration of the analogous properties of the solutions to linear second-order P.D.E.'s.

4 Boundary Value Problems

4.1 The Space of Solutions

In the last chapter we established for linear second-order O.D.E.'s that:

1. The general solution to the inhomogeneous problem is given by a particular solution plus the general solution to the homogeneous equation.

2. The space of solutions to the homogeneous equation forms a two-dimensional vector space.

3. The initial-value problem in which the value of y and y' are both specified, and do not both vanish, at a given point, x_0, guarantees a unique solution.

This section is devoted to exploring the same questions for solutions to P.D.E.'s.

The general linear second-order P.D.E. is given by

$$\mathcal{L}\psi = f \tag{4.1.1}$$

in which f is a known function of the independent variables, x, y, z and t, and the differential operator, $\mathcal{L}$, is given by eq.(2.1.1). The crucial property that this operator shares with eq.(3.2.2) is the linearity, (3.2.3), of $\mathcal{L}$. Just as was the case for the O.D.E., this ensures that the difference, $\psi_h = \psi_2 - \psi_1$, between any two solutions to eq. (4.1.1) satisfies the homogeneous equation

$$\mathcal{L}\ \psi_h = 0. \tag{4.1.2}$$

The strategy for finding the general solution to (4.1.1) is the same as for the corresponding O.D.E.. It suffices to find the general solution to the homogeneous problem, (4.1.2), together with any particular solution to (4.1.1).

The linearity of $\mathcal{L}$ as defined by eq.(2.1.1) guarantees that the set of solutions to the homogeneous equation, (4.1.2), forms a vector space. Addition of vectors and scalar multiplication is defined pointwise: $[\psi_1 + \psi_2](x, y, z, t) = \psi_1(x, y, z, t) + \psi_2(x, y, z, t)$ and $[\alpha\psi](x, y, z, t) = \alpha[\psi(x, y, z, t)]$. We again therefore expect to be able to write the general solution to (4.1.2) as a linear combination of a basis of solutions: $\psi = \sum_k c_k \psi_k$.

At this point the first significant deviation from the previous chapter appears. Whereas the space of solutions to the homogeneous O.D.E. is two-dimensional, the space of solutions to the homogeneous P.D.E. is infinite-dimensional. To see this it suffices to consider an example. Take the P.D.E. to be the two-dimensional wave equation:

$$\frac{\partial^2\psi}{\partial x^2} = \frac{\partial^2\psi}{\partial t^2}. \tag{4.1.3}$$

This is to be solved for the dependent variable, $\psi(x, t)$. This equation is simple enough to be directly solved. To do so, change variables to $u = x + t$ and $v = x - t$. In terms of these variables the (4.1.3) becomes:

$$\frac{\partial^2\psi}{\partial u \partial v} = 0. \tag{4.1.4}$$

The solution is therefore $\psi = A(x-t) + B(x+t)$ in which $A(v)$ and $B(u)$ are arbitrary *functions*. The vector space of solutions to this P.D.E. is therefore infinite dimensional. This is the major complication in extending the results of the theory of O.D.E.'s to the present case. It will prove to be possible, however, to construct a basis of solutions, $\psi_n(x,y,z,t)$ with $n = 1, 2, ...$, for the vector space of solutions to (4.1.2) subject to appropriate homogeneous boundary conditions. This is the subject of section 11. Since the sum over the index, n, that labels the basis functions includes an infinite number of terms, some care must be taken about its convergence.

4.2 Boundary-Value Problems

We turn now to the question of which boundary-value problems guarantee a unique solution for the P.D.E.'s listed in section 2. Consider for illustrative purposes Poisson's equation (2.1.3), the (inhomogeneous) diffusion equation and the (inhomogeneous) wave equation. (The inhomogeneous versions of these last two equations differ from those listed in (2.1.7) and (2.1.8) through the appearance of a known function, $f(x,y,z,t)$, on their right-hand-sides.) Each of these equations involves one more time derivative than the previous one and so requires a different type of boundary-value information.

ELLIPTIC EQUATIONS (*example: Poisson's Equation*)

Poisson's equation is an example of a *elliptic* differential equation. These are defined to be those P.D.E.'s for which the matrix of coefficients, $A_{ij}(x)$, appearing in the general form (2.1.1), has eigenvalues $\lambda_i(x)$, that nowhere vanish and all have the same sign. We wish to prove the uniqueness of the solution to the following *Dirichlet* problem:

Theorem 4.1. *Suppose $\psi(x,y,z)$ satisfies Poisson's equation, $\nabla^2\psi = f(x,y,z)$, throughout a closed, bounded region, R, together with* Dirichlet *conditions, $\psi = a(x,y,z)$ on its boundary, ∂R. a and f are known functions. Then ψ is unique.*

Proof. To prove this result, suppose that there are two distinct solutions, ψ_1 and ψ_2, to the given boundary-value problem. We prove that their difference, $u = \psi_2 - \psi_1$, vanishes, contradicting the assumption that they are distinct. The linearity of the P.D.E. implies that u satisfies the homogeneous (Laplace's) equation: $\nabla^2 u = 0$ throughout R with the boundary condition $u = 0$ on ∂R.

In order to see why these conditions imply that u vanishes throughout R consider the following manipulations that start with the vanishing of $\nabla^2 u$:

$$\begin{aligned} 0 = \int_R u\nabla^2 u \, dV &= \int_R \left[\nabla\cdot(u\nabla u) - (\nabla u)\cdot(\nabla u)\right] \, dV \\ &= \int_{\partial R} (u\nabla u)\cdot\mathbf{n} \, dS - \int_R (\nabla u)\cdot(\nabla u) \, dV \end{aligned} \tag{4.2.1}$$

The divergence theorem, (2.2.5), was used in rewriting the first term on the right-hand-side as a surface integral. The boundary information is that $u = 0$ everywhere on ∂R. This is sufficient to ensure that the boundary term above vanishes. The conclusion is then that $\int_R (\nabla u) \cdot (\nabla u)\, dV = 0$.

The integrand, $(\nabla u) \cdot (\nabla u)$, of this integral is strictly nonnegative. This implies that its integral is also nonnegative. If the integrand is sufficiently smooth the integral can vanish if and only if the integrand itself does throughout R, implying that $(\nabla u) \cdot (\nabla u) = 0$ everywhere in R. We see that ∇u must vanish everywhere in R, or equivalently, $u =$ constant in R. This last relation, together with the boundary information that $u = 0$ on ∂R implies that $u = 0$ everywhere in R as was required to be shown. We conclude that the solution to the Dirichlet problem is unique. □

Another common boundary-value problem specifies the normal derivative of the unknown function on the boundary of the region in question: *i.e.* $\mathbf{n} \cdot \nabla \psi = a(x, y, z)$ on ∂R. This is the boundary condition encountered earlier in the derivation of the diffusion equation. Such a boundary condition is known as a *Neumann* condition. The uniqueness result in this case takes the form of a:

Theorem 4.2. *Suppose $\psi(x, y, z)$ satisfies Poisson's equation,*

$$\nabla^2 \psi = f(x, y, z),$$

throughout a closed, bounded region, R, together with Neumann *conditions, $\mathbf{n} \cdot \nabla \psi = a(x, y, z)$ on its boundary, ∂R. a and f are specific known functions. Then ψ is unique up to an additive constant.*

Proof. What is to be proven in this case is that any two solutions of these conditions differ by a constant throughout all of R. That the addition of a constant to a solution produces another solution may be seen by inspection since all of the conditions involve only derivatives of u. The proof that this is the only such arbitrariness in the solution is identical to that of the previous theorem regarding Dirichlet conditions. The difference between any two solutions, $u = \psi_2 - \psi_1$, satisfies Laplace's equation and has a vanishing normal derivative on ∂R: $\mathbf{n} \cdot \nabla u = 0$. Repeating the previous argument, the boundary condition enters only in ensuring that the surface term vanishes: $\int_{\partial R} (u \nabla u) \cdot \mathbf{n}\, dS = 0$. This also follows from Neumann conditions. As before the conclusion is that ∇u vanishes throughout R, implying that u is a constant. Unlike the previous theorem concerning Dirichlet boundary conditions for u, it does *not* follow in this case that u must vanish everywhere. □

An obvious extension is to the case where Dirichlet conditions are imposed on part of the boundary and Neumann conditions are imposed on the rest. If Dirichlet conditions are chosen on *any* segment of ∂R whatsoever the solution is unique. This result may be stated as a:

Corrolary 4.1. *Consider a function, $\psi(x, y, z)$, satisfying Poisson's equation, $\nabla^2 \psi = f(x, y, z)$, throughout a closed, bounded region, R, with a piecewise-smooth boundary $\partial R = \cup_{k=1}^{n} B_k$. If, for each boundary segment B_k, ψ satisfies either* Neumann *conditions, $\mathbf{n} \cdot \nabla \psi = a_k(x, y, z)$,* *or Dirichlet conditions, $\psi = a_k(x,y,z)$ in which a_k are known functions defined for each B_k and Dirichlet conditions are chosen on at least one of the B_k, then ψ is unique.*

Proof. The proof is just as for the uniqueness of Dirichlet conditions. The boundary term in this case is:

$$\int_{\partial R} (u\nabla u)\cdot \mathbf{n}\, dS = \sum_{k=1}^{n} \int_{B_k} (u\nabla u)\cdot \mathbf{n}\, dS \tag{4.2.2}$$

which vanishes by virtue of the boundary conditions separately imposed for each segment, B_k It is of course crucial for this argument that Dirichlet conditions be imposed on at least one of the segments. This is because if Neumann conditions are imposed on all segments, then the problem is invariant with respect to constant shifts in ψ so the solution is in this case only unique up to a constant. □

A final generalization that is of physical interest is the case in which the region R is unbounded, such as when R is the exterior of a bounded region R'. In this case uniqueness can be proven subject to some extra conditions on how ψ behaves 'at infinity'. Before stating this theorem we need some definitions concerning the asymptotic behaviour of functions at infinity.

Definition 4.1. *A function, $f(x)$, is said to vanish 'faster than order x^{-p}', with $p>0$, as $x\to\infty$ if and only if the following limits are zero:*

$$\lim_{x\to\infty}[x^n f(x)] = 0 \qquad \textit{for all}\, n \le p. \tag{4.2.3}$$

This is denoted by the expression: $f(x) = o(x^{-p})$ as $x\to\infty$.

Definition 4.2. *$f(x)$ is said to vanish 'as fast as order x^{-p}', with $p>0$, as $x\to\infty$ if and only if the following limits vanish:*

$$\lim_{x\to\infty}[x^n f(x)] = 0 \qquad \textit{for all}\, n < p. \tag{4.2.4}$$

This is denoted by the expression: $f(x) = O(x^{-p})$ as $x\to\infty$. The only difference from the previous definition is the strict inequality $n<p$ in eq. (4.2.4).

Some simple properties of these definitions follow as immediate consequences of the properties of limits. Three are listed below:

$$f(x) = o(x^{-p}) \text{ and } g(x) = o(x^{-q}) \quad \text{imply} \quad f(x)+g(x) = o(x^{-r}) \tag{4.2.5}$$

with r the smaller of p and q. Also:

$$f(x) = o(x^{-p}) \text{ and } g(x) = o(x^{-q}) \quad \text{imply} \quad f(x)g(x) = o(x^{-p-q}) \tag{4.2.6}$$

$$f(x) = o(x^{-p}) \quad \text{implies} \quad \frac{df}{dx} = o(x^{-p-1}) \tag{4.2.7}$$

Corrolary 4.2. *Suppose $\psi(x,y,z)$ satisfies Poisson's equation, $\nabla^2\psi = f(x,y,z)$, throughout a closed, unbounded region, R, with boundary ∂R. Suppose also that ψ satisfies Dirichlet (or Neumann) conditions, $\psi = a(x,y,z)$ on ∂R and that $\psi = b(\theta,\phi) + o(r^{-1/2})$ as $r \to \infty$. r, θ and ϕ are the usual spherical polar coordinates and a, b and f are given functions. Then ψ is unique.*

This can be summarized as the statement that a well-posed boundary-value problem for Poisson's equation in an unbounded region is the same as that for a bounded region provided that 'infinity' is treated as an additional boundary on which asymptotic conditions must be specified.

Proof. The proof is as for the original theorem with a slight modification. Consider the closed, bounded region, $R(\rho)$, defined as the intersection of R and the interior of a large sphere, $S(\rho)$, of radius ρ. The boundary of this region is the union $\partial R(\rho) = \partial R \cup \sigma(\rho)$, where $\sigma(\rho)$ is that segment of the boundary of the sphere that intersects R. The line of argument is to repeat the original proof for $R(\rho)$ and then to take the limit $\rho \to \infty$.

Defining $u = \psi_2 - \psi_1$ for two solutions of the given boundary-value problem, we wish to argue that $u = 0$ throughout R. u satisfies Laplace's equation. Consider the vanishing quantity $0 = u\nabla^2 u$ integrated throughout $R(\rho)$. Application of the divergence theorem gives:

$$0 = \int_{\partial R} u\nabla u \cdot \mathbf{n}\, dS + \int_{\sigma(\rho)} u\nabla u \cdot \mathbf{n}\, dS - \int_{R(\rho)} (\nabla u)^2\, dV. \tag{4.2.8}$$

The integral over ∂R vanishes because of the Dirichlet or Neumann boundary conditions chosen there for ψ. The final claim to be established is that although the integral over $\sigma(\rho)$ need not vanish for finite ρ, its limit as $\rho \to \infty$ is zero. This implies that the limit of the integral over $R(\rho)$ also must vanish as $\rho \to \infty$. Since this limit is precisely the integral of $(\nabla u)^2$ over R, the remainder of the argument for a bounded region goes through, implying $u = 0$ throughout all of R.

All that remains is to show that the limit as $\rho \to \infty$ of the integral over $\sigma(\rho)$ vanishes. This follows from the assumed behaviour of ψ for large r. Together with property (4.2.5) listed above, the large-r behaviour of u is determined to be $u = o(r^{-1/2})$. Therefore, on $\sigma(\rho)$, property (4.2.7) gives $\mathbf{n} \cdot \nabla u = du/dr = o(r^{-3/2})$. Property (4.2.6) finally gives $u(\mathbf{n} \cdot \nabla u) = o(r^{-2})$. The desired limit then is:

$$\lim_{\rho\to\infty} \int_{\sigma(\rho)} (u\nabla u) \cdot \mathbf{n}\, dS = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin\theta \lim_{\rho\to\infty} \left[r^2 u \frac{du}{dr} \right]\Bigg|_{r=\rho} = 0 \tag{4.2.9}$$

The last equality follows from the definition of $o(r^{-p})$ and the fact that $u\frac{du}{dr}$ is $o(r^{-2})$. This completes the proof. $\square$

PARABOLIC EQUATIONS (*example: The Diffusion Equation*)

A *parabolic* P.D.E. is defined to be one for which the matrix of coefficients, $A_{ij}(x)$ in the general form (2.1.1) has one eigenvalue that vanishes everywhere, with all of the rest never vanishing and everywhere the same sign. The coordinate corresponding to the zero eigenvalue in a physical problem is the time coordinate, t. The zero eigenvalue means that the P.D.E. is only first-order in the time variable. The diffusion equation, (2.1.7), is an example of such a P.D.E.. From experience with the O.D.E.'s of mechanics we expect on physical grounds that an initial condition must be specified for the time variable in addition to giving all of the boundary information required for a time-independent problem. This motivates the choice of the following boundary-value problem, the uniqueness of which is now proven.

Theorem 4.3. *Suppose $\psi(x,y,z,t)$ satisfies the inhomogeneous diffusion equation,*

$$\frac{\partial\psi}{\partial t} - \kappa\nabla^2\psi = f(x,y,z),$$

throughout a closed, bounded region, R, together with Neumann conditions, $\mathbf{n}\cdot\nabla\psi = a(x,y,z,t)$ (or Dirichlet conditions, $\psi = a(x,y,z,t)$) on its boundary ∂R, and the initial condition $\psi(x,y,z,t=0) = b(x,y,z)$. The functions a , b and f are given. Then ψ is unique.

Proof. The proof starts as for the elliptic case. If ψ_1 and ψ_2 are both solutions of the given boundary-value problem, define $u = \psi_2 - \psi_1$. u satisfies the homogeneous diffusion equation throughout R and vanishes throughout R at $t = 0$. The normal derivative of u also vanishes everywhere on ∂R for all t. Consider the following function of t only:

$$U(t) = \int_R u^2\, dV. \tag{4.2.10}$$

The diffusion equation allows the derivation of an O.D.E. governing the time-evolution of U. Differentiating eq. (4.2.10) with respect to t gives:

$$\begin{aligned}\frac{dU}{dt} &= 2\int_R u\frac{\partial u}{\partial t}\, dV \\ &= 2\kappa\int_R u\nabla^2 u\, dV \\ &= 2\kappa\left[\int_{\partial R}(u\nabla u)\cdot\mathbf{n}\, dS - \int_R(\nabla u)^2\, dV\right] \\ &= -2\kappa\int_R(\nabla u)^2\, dV.\end{aligned} \tag{4.2.11}$$

In the last equality the boundary information that $\mathbf{n}\cdot\nabla u$ (or, for Dirichlet conditions, u) vanishes has been used to kill off the surface integral.

Eq. (4.2.11) implies that the derivative of U is strictly nonpositive. We also know from the definition (4.2.10) that U itself is strictly nonnegative. Furthermore, at the initial time $t = 0$ the initial condition states that $u(x,y,z,t=0) = 0$, which implies that $U(0) = 0$. The only nonnegative, decreasing function of t that starts at zero at $t = 0$ is the function $U(t) = 0$ for all $t > 0$. Since the integrand in eq. (4.2.10) is nonnegative and integrates to zero for all $t > 0$ it follows that $u = 0$ throughout R for all times, as was required to be shown. □

Corrolary 4.3. *The corollaries derived earlier for Poisson's equation also follow here from the identical reasoning.*

HYPERBOLIC EQUATIONS (*example: The Wave Equation*)

An *hyperbolic* P.D.E. is one for which the matrix of coefficients, $A_{ij}(x)$ in the general form (2.1.1) has eigenvalues that nowhere vanish and for which one eigenvalue differs in sign from all of the rest. The coordinate with the odd-sign eigenvalue is again the time coordinate, t. The P.D.E. is in this case second-order in both the time and space variables. The standard representative of such a P.D.E. is the wave equation, (2.1.8). From experience with O.D.E.'s we expect that both the value of the unknown function and its time derivative must be specified at the initial time, in addition to giving all of the spatial boundary information. This corresponds to choosing the initial 'position' and 'velocity' for the problem. The resulting well-posed boundary-value problem is easily proven to be unique.

Theorem 4.4. *Suppose $\psi(x, y, z, t)$ satisfies the inhomogeneous wave equation,*

$$-\frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} + \nabla^2\psi = f(x, y, z),$$

throughout a closed, bounded region, R, together with Neumann conditions, $\mathbf{n}\cdot\nabla\psi = a(x, y, z, t)$ (or Dirichlet conditions $\psi = a$) on its boundary ∂R, and the initial conditions $\psi(x, y, z, t = 0) = b(x, y, z)$ and with both $\left.\frac{\partial\psi}{\partial t}\right|_{t=0} = c(x, y, z)$ specified at $t = 0$. a, b, c and f are known functions. Then ψ is unique.

Proof. The proof needs only minor modifications from that used for the diffusion equation. The game as usual is to prove that the difference, $u = \psi_2 - \psi_1$, between any two solutions of the given boundary-value problem must vanish. u satisfies in this case the wave equation with the initial conditions that it, and its first time derivative, vanish throughout R at $t = 0$. Either u itself or its normal derivative also vanishes on ∂R for all t.

Define the following two nonnegative functions of t only:

$$\begin{aligned} V(t) &= \int_R \left(\frac{\partial u}{\partial t}\right)^2 dV \\ W(t) &= \int_R (\nabla u)^2 \, dV. \end{aligned} \tag{4.2.12}$$

Their time-evolution is related by the wave equation:

$$\begin{aligned}
\frac{dV}{dt} &= 2\int_R \frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2}\, dV \\
&= 2v^2 \int_R \frac{\partial u}{\partial t}\nabla^2 u\, dV \\
&= 2v^2 \left[\int_{\partial R} \frac{\partial u}{\partial t}\mathbf{n}\cdot\nabla u\, dS - \int_R \left(\frac{\partial}{\partial t}\nabla u\right)\cdot\nabla u\, dV\right] \\
&= 2v^2 \int_{\partial R} \frac{\partial u}{\partial t}\mathbf{n}\cdot\nabla u\, dS - v^2\frac{dW}{dt}.
\end{aligned} \tag{4.2.13}$$

The boundary condition $\mathbf{n}\cdot\nabla u = 0$ guarantees the vanishing of the surface integral. (For Dirichlet conditions the condition $u = 0$ throughout ∂R for all t implies, after differentiation with respect to t, that $\frac{\partial u}{\partial t} = 0$ throughout R, similarly making the surface term zero.) Eq. (4.2.14) is a then simple O.D.E. to integrate, and states that $V + v^2 W$ is independent of t. The initial condition that fixes the integration constant is $V(0) = W(0) = 0$, which follows from the vanishing of u and $\frac{\partial u}{\partial t}$ throughout R at $t = 0$. The solution is therefore $V(t) = -v^2 W(t)$ for all t. Since $V(t)$ and $W(t)$ are both strictly nonnegative it follows that they must both vanish for all t. The vanishing of W implies that $\nabla u = 0$ within R for all time and this, with the initial condition, forces u to be zero throughout R for all t, as was required. □

The extensions to unbounded R and piecewise continuous ∂R follow as before.

5 Separation of Variables

The previous chapters have been devoted to general properties of the solutions of the O.D.E.'s and P.D.E.'s that appear in mathematical physics. The present chapter, on the other hand, turns to the problem of actually constructing these solutions. The method to be outlined unfortunately does not furnish the general solution to an arbitrary second-order linear P.D.E.. It does, however, allow the construction of general solutions to many of the boundary-value problems that commonly occur.

The approach is to look for solutions to the differential equation that have a product form, $\psi(x,y,z,t) = X(x)Y(y)Z(z)T(t)$ for example, with the intention of reducing the partial differential equation in question to a system of ordinary differential equations that can be solved by general techniques. (These techniques are themselves the topics of later chapters.) Some key questions are: (i) under what circumstances do such solutions exist and (ii) what relation do they bear on the general solution to the given problem. The utility of this method of solution obviously relies on there being some nontrivial set of problems whose solution can be found using these techniques. A partial answer to these questions is given later in this chapter following a more detailed description of the method.

5.1 An Example and Some General Features

Consider the following boundary-value problem: Solve the diffusion equation,

$$\frac{\partial \psi}{\partial t} = \kappa \nabla^2 \psi \tag{5.1.1}$$

within the cubic region $0 \le x \le L$, $0 \le y \le L$ and $0 \le z \le L$ subject to the boundary conditions that ψ vanishes on the surfaces $x = 0$, $x = L$, $y = 0$ and $y = L$. On the remainder of the boundary, $z = 0$ and $z = L$, choose Neumann conditions: $\frac{\partial \psi}{\partial z} = 0$. The initial condition is chosen to be: $\psi(t = 0) = A \sin(\pi x/L) \sin(\pi y/L) \cos(\pi z/L)$.

(Notice that the given initial condition is consistent with the boundary conditions at $t = 0$. This is obviously a necessary condition for the existence of a solution.)

As mentioned in the introduction we search for a solution of the product form:

$$\psi(x,y,z,t) = X(x)Y(y)Z(z)T(t). \tag{5.1.2}$$

Substituting this into the differential equation (5.1.1) and dividing the result through by ψ gives:

$$\frac{1}{T}\frac{dT}{dt} = \kappa \left[\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} \right] \tag{5.1.3}$$

Now comes the key argument. The main point is that eq. (5.1.3) must hold for all (x, y, z) in the cube and for all $t > 0$. Since each term in the equation depends on a different coordinate the only way that it can be satisfied for all values of the coordinates is for each term to be separately constant. That is, eq. (5.1.3) is equivalent to the following set of *ordinary* differential equations:

$$
\begin{aligned}
\frac{dT}{dt} &= c_1 T \\
\frac{d^2X}{dx^2} &= c_2 X \\
\frac{d^2Y}{dy^2} &= c_3 Y \\
\frac{d^2Z}{dz^2} &= c_4 Z \\
\text{with} \qquad c_1 &= \kappa(c_2 + c_3 + c_4).
\end{aligned} \tag{5.1.4}
$$

These are easily solved:

$$
T(t) = C_1 \exp(c_1 t) \tag{5.1.5}
$$

and:

$$
\begin{aligned}
X(x) &= C_2 \exp(\sqrt{c_2}x) + C_3 \exp(-\sqrt{c_2}x) \qquad \text{if } c_2 > 0 \\
&= C_2 \sin(\sqrt{-c_2}x) + C_3 \cos(\sqrt{-c_2}x) \qquad \text{if } c_2 < 0 \\
&= C_2 + C_3 x \qquad \text{if } c_2 = 0
\end{aligned} \tag{5.1.6}
$$

with similar solutions for $Y(y)$ and $Z(z)$.

A key feature of the given problem is that the spatial boundary conditions are to be imposed on surfaces defined by equations of the form 'coordinate=constant'. This, together with the form of the boundary condition taken there, allows the boundary information to be enforced separately for the functions X, Y and Z. The integration constants become partially determined in this way giving:

$$
\begin{aligned}
X_l(x) &= C_2' \sin\left(\frac{\pi l x}{L}\right), \qquad l = 1, 2, ... \\
Y_m(y) &= C_3' \sin\left(\frac{\pi m y}{L}\right), \qquad m = 1, 2, ... \\
Z_n(z) &= C_4' \cos\left(\frac{\pi n z}{L}\right), \qquad n = 0, 1, ... \\
T_{lmn}(t) &= C_1 \exp(c_1 t) \\
\text{with} \qquad c_1 &= -\kappa \frac{\pi^2}{L^2}(l^2 + m^2 + n^2)
\end{aligned} \tag{5.1.7}
$$

Notice that the solutions are not unique, being labelled by the positive (or, for n, nonnegative) integers l, m and n. This is because the boundary-value problems given by the O.D.E.'s (5.1.4) for X, Y and Z, *together with their boundary conditions*, are linear and homogeneous. Any linear combination of solutions is therefore also a solution and the space of such solutions forms a vector space in the sense discussed in chapters 3. The corresponding solution space to the full P.D.E. together with its spatial boundary conditions is therefore also a vector space. The temporal boundary condition, on the other hand, is not homogeneous and so it is this piece of information that picks out the unique element of the vector space that satisfies the full problem. We see therefore that a very large class of solutions to the P.D.E. (5.1.1) is given by the general linear combination:

$$\psi(x,y,z,t) = \sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=0}^{\infty} C_{lmn} X_l(x) Y_m(y) Z_n(z) T_{lmn}(t) \tag{5.1.8}$$

in which the constant coefficients, C_{lmn}, are constrained only by the requirement that this sum converge. The particular choice of these constants that satisfies the initial condition given in this example is $C_{111} = A$, with all others zero. That is, the unique solution to the given boundary-value problem is the particular combination:

$$\psi(x,y,z,t) = A \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \cos\left(\frac{\pi z}{L}\right) \exp\left(-3\kappa \frac{\pi^2}{L^2} t\right) \tag{5.1.9}$$

This example illustrates the general features of the approach, as well as some of its limitations. The basic idea is to convert the P.D.E. (5.1.1) into a set of O.D.E.'s, like eqs. (5.1.4), which can then be solved. In itself this would appear to be only small progress since only the most simple and contrived problems have solutions satisfying the separated ansatz (5.1.2). The real utility of the method lies in the ability to expand a general solution of a homogeneous boundary-value problem in terms of a basis of functions that satisfy this ansatz.

The big question becomes: to which problems does this method apply? From the example we can see there are several necessary conditions. They are:

1. The region, R, in which the boundary-value problem is posed must have boundaries, ∂R, that can be expressed as surfaces of constant coordinate values, *i.e.*by equations of the form 'coordinate=constant'. This in itself is a minor requirement since the coordinates used can generally be chosen to be constant on a given boundary.

2. The P.D.E. must also be *separable* in the coordinates defined by requirement (1). This means that once the product ansatz (5.1.2) is chosen for these coordinates, the differential equation can be made to become a sum of terms each depending on a single coordinate. This is crucial if the P.D.E. is to be reduced to a set of equivalent O.D.E.'s.

3. The boundary conditions satisfied on ∂R must be such that they may be imposed separately on the component functions of each coordinate.

4. Finally 'most' of the boundary conditions should, in practice, be linear and homogeneous so that the space of solutions forms a vector space. It is then necessary to show that any element of this space may be expanded in terms of a basis of separated functions. An explicit expression for the expansion coefficients corresponding to a given function must then also be known in order to use the complete boundary-value information to solve for the unknown function.

Property (1) can in principle be satisfied by an artful choice of coordinates given the geometry of a specific problem. In practice, the geometries that often arise are regions that are rectangular, cylindrical or spherical. This reflects the symmetries under translations or rotations that underly the physics of the problem. We therefore focus explicitly on separation of variables in cylindrical and spherical coordinates in the remainder of this chapter.

Property (2) is the key one that may or may not be satisfied by the problem of interest. Given that it is satisfied, the demonstration of remaining properties (iii) and (iv) may be done fairly generally and is the subject of *Sturm-Liouville* theory in chapter 11.

We turn now to separation of variables in cylindrical and spherical coordinates. Apart from serving as useful examples of the method, this excercise introduces some O.D.E.'s that arise particularly frequently, and so whose solutions are to be examined in some detail in later chapters.

5.2 Cylindrical Coordinates

Cylindrical coordinates are most convenient for problems that are symmetric under rotations in a plane and under translations in the direction perpendicular to that plane. The coordinates (ρ, ξ, z, t) are related to rectangular coordinates by:

$$\begin{aligned} x &= \rho \cos \xi, \qquad 0 \le \xi \le 2\pi \\ y &= \rho \sin \xi, \qquad 0 \le \rho < \infty \\ z &= z \\ t &= t \end{aligned} \tag{5.2.1}$$

For the angular coordinate, $\xi = 0$ labels the same point as $\xi = 2\pi$ and so any smooth single-valued function of (x, y, z, t) must be *periodic* in ξ with period 2π. That is, it must be invariant under the shift $\xi \to \xi + 2\pi$.

We again take the diffusion equation (2.1.7) as our example. In cylindrical coordinates the Laplacian operator, ∇^2, defined by eq. (2.1.2) becomes:

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \psi}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 \psi}{\partial \xi^2} + \frac{\partial^2 \psi}{\partial z^2}. \tag{5.2.2}$$

so the diffusion equation is:

$$\frac{\partial \psi}{\partial t} - \kappa \left[\frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \psi}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 \psi}{\partial \xi^2} + \frac{\partial^2 \psi}{\partial z^2} \right] = 0. \tag{5.2.3}$$

The separation *ansatz* in this case is:

$$\psi(\rho, \xi, z, t) = R(\rho)\Xi(\xi)Z(z)T(t). \tag{5.2.4}$$

Substitution of (5.2.4) into (5.2.3) and division by ψ gives:

$$\frac{1}{T}\frac{dT}{dt} - \kappa \left[\frac{1}{R}\frac{d^2 R}{d\rho^2} + \frac{1}{\rho R}\frac{dR}{d\rho} + \frac{1}{\rho^2 \Xi}\frac{d^2 \Xi}{d\xi^2} + \frac{1}{Z}\frac{d^2 Z}{dz^2} \right] = 0. \tag{5.2.5}$$

The P.D.E. has become of the form: $f(t)+g(\rho,\xi)+h(z)=0$ so the z- and t- dependence can be separated as before. The t- and z-dependent parts of the equation must be constants:

$$\begin{aligned}\frac{dT}{dt} &= c_1 T(t)\\ \frac{d^2Z}{dz^2} &= c_2 Z(z)\end{aligned} \tag{5.2.6}$$

$$\frac{1}{R}\frac{d^2R}{d\rho^2}+\frac{1}{\rho R}\frac{dR}{d\rho}+\frac{1}{\rho^2\Xi}\frac{d^2\Xi}{d\xi^2} = \frac{c_1}{\kappa}-c_2$$

The last of these equations may be further separated if it can be put into the form $f(\rho)+g(\xi)=0$. To this end multiply through by ρ^2. The result has the desired form and gives the additional separated equations:

$$\begin{aligned}\frac{d^2\Xi}{d\xi^2} &= c_3\Xi\\ \frac{d^2R}{d\rho^2}+\frac{1}{\rho}\frac{dR}{d\rho}+\frac{c_3}{\rho^2}R+\left(c_2-\frac{c_1}{\kappa}\right)R&=0.\end{aligned} \tag{5.2.7}$$

The problem remains to solve these O.D.E.'s. Those for $T(t)$, $Z(z)$ and $\Xi(\xi)$ are elementary and have solutions given above in eqs. (5.1.5) or (5.1.6). Furthermore, the very definition of the coordinate ξ implies a linear and homogeneous boundary condition for Ξ: $\Xi(\xi+2\pi)=\Xi(\xi)$. This partially determines the separation constant c_3 and so gives the following family of solutions:

$$\begin{aligned}\Xi_n(\xi) &= C_1\cos(n\xi)+C_2\sin(n\xi)\\ c_3 &= -n^2, \qquad n=0,\pm1,\pm2,...\end{aligned} \tag{5.2.8}$$

The remaining equation for $R(\rho)$ is more difficult and the properties of its solutions are explored in some detail in the remaining chapters. It is conventional to rewrite the O.D.E. for $R(\rho)$ in a more conventional form. To do so consider $\alpha^2=|c_2-c_1/\kappa|$. Changing the independent variable to $x=\alpha\rho$ gives:

$$\frac{d^2R}{dx^2}+\frac{1}{x}\frac{dR}{dx}+\left(\pm1-\frac{n^2}{x^2}\right)R=0. \tag{5.2.9}$$

The constant α has dropped right out of the equation. The sign $\pm$ denotes the sign of c_2-c_1/κ. When the sign in (5.2.9) is positive this is called *Bessel's equation*. With the negative sign it is known as Bessel's *modified* equation. Solutions to the modified equation are obtained from solutions to Bessel's equation by simply evaluating the result at an imaginary argument.

5.3 Spherical Polar Coordinates

As might be expected spherical coordinates are appropriate to problems with spherical symmetry. These coordinates are denoted (r, θ, ϕ, t) and are defined by:

$$
\begin{aligned}
x &= r\cos\phi\sin\theta \\
y &= r\sin\phi\sin\theta \\
z &= r\cos\theta \\
t &= t.
\end{aligned}
\tag{5.3.1}
$$

The coordinates run over the range $0 < \theta < \pi$, $0 \le \phi \le 2\pi$ and $0 < r < \infty$. The angle ϕ is periodic in the sense that $\phi = 0$ and $\phi = 2\pi$ are identified. Any single-valued smooth function must therefore be periodic under $\phi \to \phi + 2\pi$.

We sketch the procedure using the diffusion equation as the illustrative example. In these coordinates the Laplacian is:

$$
\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}. \tag{5.3.2}
$$

allowing the diffusion equation to be written:

$$
\frac{\partial\psi}{\partial t} = \kappa\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right]. \tag{5.3.3}
$$

The separation ansatz is:

$$
\psi(r, \theta, \phi, t) = R(r)\Theta(\theta)\Phi(\phi)T(t). \tag{5.3.4}
$$

Using this in eq. (5.3.3) and dividing through by ψ gives, for T, the O.D.E. of eq. (5.1.4) together with the following equation for R, Θ and Φ:

$$
\frac{1}{Rr^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{1}{\Theta r^2\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) + \frac{1}{\Phi r^2\sin^2\theta}\frac{d^2\Phi}{d\phi^2} = \frac{c_1}{\kappa}. \tag{5.3.5}
$$

Separate the remaining variables one by one. Multiplying eq. (5.3.5) through by $r^2\sin^2\theta$ allows the ϕ-dependence to be separated:

$$
\frac{d^2\Phi}{d\phi^2} = c_2\Phi(\phi) \tag{5.3.6}
$$

and

$$
\frac{\sin^2\theta}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{\sin\theta}{\Theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) + c_2 - \frac{c_1}{\kappa}r^2\sin^2\theta = 0 \tag{5.3.7}
$$

Finally, dividing by $\sin^2\theta$ separates the remaining two variables.

$$
\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \left(c_3 - \frac{c_1}{\kappa}r^2\right)R = 0 \tag{5.3.8}
$$

$$
\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) + \left(\frac{c_2}{\sin^2\theta} - c_3\right)\Theta = 0. \tag{5.3.9}
$$

The O.D.E.'s for $T(t)$, and $\Phi(\phi)$ are easily solved with solutions identical to those of eqs. (5.1.5) and, after using the boundary condition for ϕ, (5.2.8). The separation constant, c_2 is determined to be $c_2 = -m^2$ for $m = 0, \pm 1, ...$

The solutions to eqs. (5.3.8) and (5.3.9) are more difficult to determine. We content ourselves here to putting the remaining O.D.E.'s into canonical form. Their solutions are found in subsequent chapters. For the equation governing Θ perform the change of variables: $x = \cos\theta$. The range of x of physical interest is therefore $-1 < x < 1$. In the new variables the O.D.E. becomes:

$$\frac{d}{dx}\left[(1-x^2)\frac{d\Theta}{dx}\right] - \left(c_3 + \frac{m^2}{1-x^2}\right)\Theta = 0. \tag{5.3.10}$$

This is called the *Associated Legendre equation*. The special case when $m = 0$ is *Legendre's equation*. The solution to (5.3.10) is to be found on the interval $x \in (-1, 1)$. Physically, the endpoints $x = \pm 1$ correspond to the lines $\theta = 0$ and $\theta = \pi$ which label the z-axis. Since there is nothing special about the z-axis we adopt the (homogeneous) boundary condition that Θ not diverge even when x approaches ± 1. As we shall see in chapter 10, it turns out that the generic solution to (5.3.10) does diverge at one of the endpoints $x = \pm 1$ unless c_3 takes special values. These values are: $c_3 = -l(l+1)$ with $l = |m|, |m|+1, ...$ This result for c_3 is used for the remainder of this section.

The radial equation for $R(r)$ is related to Bessel's equation (5.2.9). To make this connection define the new independent variable $x = \alpha r$ where α is defined by $\alpha = \sqrt{|c_1/\kappa|}$. The radial O.D.E. is then:

$$\frac{1}{x^2}\frac{d}{dx}\left(x^2\frac{dR}{dx}\right) + \left(\pm 1 - \frac{l(l+1)}{x^2}\right)R = 0. \tag{5.3.11}$$

The upper (lower) sign ± 1 in this equation corresponds to the case where $c_1 < 0$ ($c_1 > 0$). The boundary conditions for $R(r)$ usually determine this to be negative. We therefore take the upper sign in the remainder of this chapter. Finally, performing the change of dependent variable: $y(x) = R(x)/\sqrt{x}$ gives:

$$\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{(l+\frac{1}{2})^2}{x^2}\right)y = 0 \tag{5.3.12}$$

which is recognized as Bessel's equation with $n = l + \frac{1}{2}$. Solutions to this equation are often called *spherical Bessel functions* for obvious reasons.

6 Complex Variables and Analytic Continuation

6.1 Introduction

The method of separation of variables has reduced the problem of constructing solutions to the P.D.E.'s of interest to the problem of solving a set of second order, linear O.D.E.'s, of the form:

$$\frac{d^2y}{dz^2} + p(z)\frac{dy}{dz} + q(z)y = f(z). \tag{6.1.1}$$

What remains is to solve these O.D.E.'s. The method that we shall rely on is the construction of solutions by power series expansion. It is therefore crucial to understand under which circumstances such a power-series solution exists. Since the existence of a series expansion of a function is related to the analytic properties of that function, about which the theory of complex variables has much to say, we are led to formulate the differential equation with both the dependent variable, y, and the independent one, z, taking complex values.

The two goals of the following sections can be simply stated. The first is to relate the analytic properties of the solutions, $y(z)$, of eq.(6.1.1) to the corresponding properties of the coefficient functions, $p(z)$ and $q(z)$. Next, the type of O.D.E.'s that admit a series solution must be classified and solved. In either case we first require a quick reminder of the properties of functions of a single complex variable.

6.2 Review of the Calculus of a single Complex Variable

For convenience, some pertinent facts from the theory of the calculus of a single complex variable are listed here.

Any complex function of a complex variable, $z = x + iy$, can be written as $f(z, z^*) = u(x, y) + iv(x, y)$ and so is equivalent to a pair of real functions of the two real variables x and y. f is differentiable, say, with respect to z and z^* if and only if u and v are differentiable with respect to x and y.

A special role is played by functions of the form $f = f(z)$, *i.e.* those that are independent of z^*. Independence of z^* can be expressed in terms of u and v by examining the real and imaginary parts of the condition $\partial f/\partial z^* = 0$:

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = 0 \tag{6.2.1}$$

These are called the *Cauchy-Riemann* equations.

Definition 6.1. *Analytic function: A function, $f(z)$, is by definition* analytic *at a point, z_0, if it is differentiable on a neighbourhood that includes z_0. That is, the limit:*

$$\lim_{z \to z'} \frac{f(z) - f(z')}{z - z'} \tag{6.2.2}$$

exists, for z' sufficiently close to z_0.

Definition 6.2. *Singular Point: If $f(z)$ is not analytic at $z = z_0$, then z_0 is called a singular point of $f(z)$. It is an* isolated singularity *if z_0 lies within a disc, $D_c = \{z \ni |z - z_0| < c\}$, on which $f(z)$ is analytic everywhere except for the point $z = z_0$.*

Definition 6.3. *Taylor Series: If $f(z)$ is analytic at $z = z_0$, then there exists a disc, D_c, centered on z_0, in which its* Taylor series *is unique and converges uniformly to $f(z)$:*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \qquad \forall z \in D_c. \tag{6.2.3}$$

The coefficients, a_n, in this series are found by evaluating successive derivatives at $z = z_0$:

$$a_n = \frac{1}{n!} \left. \frac{d^n f}{dz^n} \right|_{z=z_0}. \tag{6.2.4}$$

Definition 6.4. *Laurent Series: If $f(z)$ is analytic throughout an annular region, $A_{a,b} = \{z \ni a < |z - z_0| < b\}$, surrounding the point $z = z_0$, then its* Laurent series *is unique and converges uniformly to $f(z)$ throughout $A_{a,b}$:*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}. \tag{6.2.5}$$

In particular, the Laurent series exists on a disc surrounding any isolated singularity.

Definition 6.5. *Poles: An isolated singularity is called a pole of order n, for n a positive integer, if in eq. (6.2.5) $b_n \neq 0$ and $b_k = 0$ for all $k > n$. A pole of order 1 is called a simple pole.*

Definition 6.6. *Essential Singularity: An isolated singularity is called an* essential singularity *if the b_k-series doesn't terminate. That is to say, for any positive integer, n, there is a $k > n$ such that $b_k \neq 0$.*

Definition 6.7. *Entire Functions: A function, $f(z)$, that is analytic for all finite z is called an entire function. The Taylor expansion of an entire function about any finite point, z_0, exists for all finite z.*

EXAMPLES:

a) $f(z) = 1/z$ has a pole of order 1 at $z = 0$ and is analytic everywhere else.

b) $f(z) = \exp(1/z) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{z}\right)^n$ has an essential singularity at $z = 0$ and is analytic elsewhere.

c) $f(z) = \log z$ is singular at $z = 0$ but the singularity is not isolated since its imaginary part, $\text{Im}\ \log z = \arg z \in (-\pi, \pi)$, is discontinuous along the negative real axis. $z = 0$ is called the *branch point* and the line of discontinuity is called the *branch cut* of $f(z)$. No Laurent expansion exists for $\log z$ about $z = 0$.

d) $f(z) = z^s$ has an isolated singularity at $z = 0$ (a pole of order $-s$) when s is a negative integer. If s is a nonnegative integer then z^s is analytic for all finite z. If s is not an integer then $z = 0$ is a singular point of z^s that is not isolated since z^s is not periodic in $\arg z$. No Laurent series exists about the branch point at $z = 0$, when s is not an integer.

e) $f(z) = e^z$ is an entire function.

Definition 6.8. *Asymptotic Forms: A function, $f(z)$, is bounded if there exists a positive real number, M such that $|f(z)| < M$ for all finite z. Similarly, a function, $f(z)$, is said to grow like z^n as $z \to \infty$, with n a positive integer, if $|f(z)| < M|z|^n$ for all finite z.*

Theorem 6.1. *Liouville's Theorem: The only entire, bounded function is the constant function: $f(z) =$ constant.*

Theorem 6.2. *The only entire function that grows like z^n as $z \to \infty$ is an n-th degree polynomial: $f(z) = \sum_{k=0}^{n} a_k z^k$.*

Definition 6.9. *Contour Integration: Suppose that $f(z)$ is a complex function and C denotes the path (or contour) in the complex plane that runs from $z = z_0$ to $z = z_1$ given by the equations: $z = c(t) = x(t) + iy(t)$, $0 \le t \le 1$ with $c(0) = z_0$ and $c(1) = z_1$. The complex contour integral, $I = \int_C f(z)\, dz$, is defined by the real integral:*

$$\begin{aligned} I &= \int_0^1 f[z(t)] \frac{dz}{dt}\, dt \\ &= \int_0^1 \left(u[x(t), y(t)] \frac{dx}{dt} - v[x(t), y(t)] \frac{dy}{dt} \right) dt \\ &\qquad + i \int_0^1 \left(u[x(t), y(t)] \frac{dy}{dt} + v[x(t), y(t)] \frac{dx}{dt} \right) dt. \end{aligned} \tag{6.2.6}$$

Theorem 6.3. *Cauchy-Goursat Theorem: Suppose C denotes a closed curve given by the equation: $z = c(t)$, $0 \le t \le 1$ with $c(0) = c(1)$. If $f(z)$ is analytic at all points on and within C then:*

$$\int_C f(z)\, dz = 0. \tag{6.2.7}$$

Notice that this implies that the contour integral along any two curves, C_1 and C_2, joining the points z_0 and z_1 in the complex plane must agree provided that the region between the two contours doesn't contain a singular point of the integrand. To see this notice that traversing contour C_1 from $t = 0$ to $t = 1$ and then traversing C_2 backwards from $t = 1$ to $t = 0$ amounts to travelling once around a closed contour, C. Therefore:

$$\int_{C_1} f(z)\, dz - \int_{C_2} f(z)\, dz = \int_C f(z)\, dz = 0 \tag{6.2.8}$$

A different way of expressing the same conclusion is that a complex integral is unchanged as the contour of integration is 'deformed' all around the complex plane provided only that the deformations never take the contour across a singular point of the integrand.

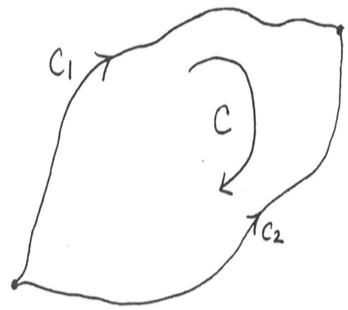

Figure 6.1. The Contours C_1, C_2 and C

Theorem 6.4. *Morera's Theorem: This is the converse to Cauchy's theorem. If* $\int_C f(z)\,dz = 0$ *for all closed curves,* C*, in a connected open domain,* D*, then* $f(z)$ *is analytic throughout* D*.*

Definition 6.10. *Residues: If* $z = z_0$ *is an isolated singularity of an analytic function,* $f(z)$*, with Laurent expansion given by eq.(6.2.5), then the residue of* $f(z)$ *at* z_0 *is defined as the value of the coefficient* b_1*. Notice that if* $f(z)$ *has a simple pole at* z_0 *then inspection of the Laurent series shows that the residue is given by:*

$$B = \lim_{z\to z_0}[(z-z_0)f(z)] = \lim_{z\to z_0}\left[\sum_{n=0}^{\infty} a_n(z-z_0)^{n+1} + b_1\right] = b_1. \tag{6.2.9}$$

Theorem 6.5. *Residue Theorem: If* C *is a closed contour within and on which a function,* $f(z)$*, is analytic except at a finite number of isolated singular points,* $\{z_1, z_2, ..., z_n\}$ *in the interior of* C*, then:*

$$\int_C f(z)\,dz = 2\pi i(B_1 + B_2 + ... + B_n) \tag{6.2.10}$$

in which the integration traverses C *in the counterclockwise sense and* B_k *denotes the residue of* $f(z)$ *at the point* z_k*.*

EXAMPLES:

a) Suppose $f(z) = 1/z$ is integrated about the closed contour $C:\ c(t) = \exp(2\pi i t)$ for $0 \le t \le 1$. This equation describes the unit circle traversed once counterclockwise starting from $z_0 = c(0) = 1$. The integral is $\int_C (1/z)\,dz = 2\pi i \int_0^1 dt = 2\pi i$. This agrees with the residue theorem since $f(z) = 1/z$ has a simple pole in the interior of C at $z = 0$ with residue $B = 1$.

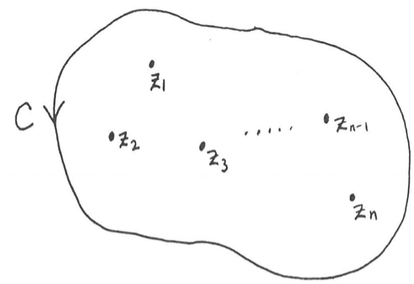

Figure 6.2. The Contour C

b) If $f(z) = 1/z^2$ is integrated over the same contour as in the previous example, the result is $\int_C (1/z^2)\ dz = 2\pi i \int_0^1 e^{-2\pi i t}\ dt = 0$. This also agrees with the residue theorem because $f(z) = 1/z^2$ has an isolated singular point at $z = 0$ but with residue $B = 0$.

c) This final example gives a contour integral representation for a series. It provides a powerful method for evaluating series. Consider the series, $S = \sum_{n=-\infty}^{\infty} a_n$. Suppose there exists a function, $f(z)$, that is analytic near the real axis and which satisfies $f(z = n) = a_n$. An integral representation of the series, S, is furnished by:

$$S = \frac{1}{2i} \int_C f(z) \cot(\pi z)\ dz \tag{6.2.11}$$

in which the contour, C, runs from $-\infty$ to ∞ just below the real axis and returns from ∞ to $-\infty$ just above the real axis. The interior of this contour is a narrow strip centered on the real axis.

To establish the validity of eq.(6.2.11) we evaluate the integral along C by residues. The only poles of the integrand are, by assumption, those of the cotangent factor and these consist of simple poles for integer $z = n$. The residue of the integrand at these poles is (using l'Hôpital's rule and point 14 above):

$$B_n = \lim_{z \to n} \frac{(z-n) f(z)}{\tan(\pi z)} = \lim_{z \to n} \frac{f(z)}{\pi \sec^2(\pi z)} = \frac{f(n)}{\pi}. \tag{6.2.12}$$

so the residue theorem gives:

$$\frac{1}{2i} \int_C f(z) \cot(\pi z)\ dz = \pi \sum_{n=-\infty}^{\infty} B_n = \sum_{n=-\infty}^{\infty} f(n) = S. \tag{6.2.13}$$

as required.

This completes the brief summary of the calculus of a complex variable.

6.3 Analytic Continuation: Uniqueness

It will rarely be possible to express the solutions to the O.D.E.'s of interest in closed form. They will instead be exhibited as a power series or in an integral representation. Such representations are generally well-defined only for a restricted region of the complex plane, R say, in which the sum or integral converges. It is in order to extend this kind of representation to more general regions that the notion of analytic continuation is required.

Statement of the problem:

Suppose a function is defined via an integral representation. For example:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt \tag{6.3.1}$$

defines a function, $\Gamma(z)$, within the region of convergence of the integral: $\mathrm{Re}\ z > 0$. This function is analytic within this region since

$$\frac{d\Gamma}{dz} = \int_0^\infty t^{z-1} e^{-t} \log t\, dt \tag{6.3.2}$$

also converges for $\mathrm{Re}\ z > 0$. The problem is: can such a definition of a function be extended in a reasonable (unique) way to the entire complex plane? Although in general there is no such natural extension, it is possible in many cases of interest. The main result takes the form of the following:

Theorem 6.6. *If two functions, $f_1(z)$ and $f_2(z)$, are analytic in a connected open region, D_1, of the complex plane, and if $f_1 = f_2$ for all z in a connected open region $D_2 \subset D_1$, then $f_1 = f_2$ throughout all of D_1.*

This states that if an analytic function is known on a complex region, D_2, and if it can be extended to an analytic function throughout a larger region, D_1, that includes D_2, then that extension is unique. The extension of an analytic function from D_2 to D_1 is known as an *analytic continuation*.

Proof. (Rather, Sketch of Proof):

To get a flavour for the argument, we present a sketch of the proof of the theorem. Consider the function $g(z) = f_1(z) - f_2(z)$. By assumption this function is analytic throughout the larger region, D_1, and vanishes throughout D_2. The argument is to show that these conditions imply that $g(z)$ vanishes throughout D_1. To show this choose any point, ζ, that lies within D_1 but not D_2, and choose a curve, C, that connects ζ to some point, z_0, within D_2. By assumption $g(z)$ is analytic along some strip of width $2w$ contained within D_1 and centered on the curve C. This implies that $g(z)$ admits a Taylor series expansion about any point of C whose radius of convergence is at least w.

The idea is now to consider a sequence of n overlapping discs of radius w and centered at points, z_k, that lie on C. We require that the first point be $z_0 \in D_2$, that the last point be

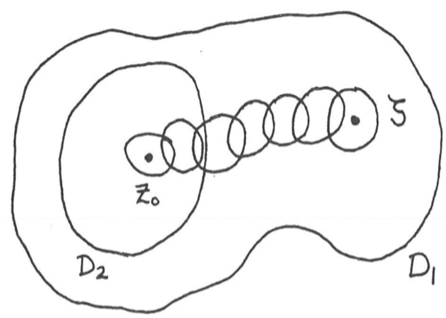

Figure 6.3. The n Overlapping Discs

$z_{n-1} = \zeta$ and that each point, z_k, lies within the neighbouring two discs. Since $g(z)$ vanishes throughout D_2 it vanishes within the first disc and so its Taylor series there is zero. This implies that $g(z)$ vanishes on the overlap with the next disc and so its power series there must also be zero. Since this series is unique, it and $g(z)$ must therefore vanish throughout the entire disc centered at z_1. Proceeding in an identical fashion implies that $g(z)$ vanishes within all of the discs and so also at ζ. Repeating this argument for any point, ζ, in D_1 implies that $g(z)$ vanishes throughout D_1 as required. $\square$

Before moving on to the construction of analytic continuations of given functions we consider a few examples that cannot be analytically continued.

1. $f(z) = \log z = \log|z| + i \arg z$ for $D = \{z \ni -\pi < \arg z < \pi\}$. If an analytic continuation to the entire plane were to exist it would have to be given by this same expression since it is unique and must agree with $f(z)$ everywhere in D. Since this expression is not continuous across the negative real axis, an analytic continuation cannot exist.

2. Consider the function defined for Im $z > 0$ by:

$$f(z) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty}{}' \frac{e^{-(m^2+n^2)}}{mz - n}. \tag{6.3.3}$$

The prime denotes the omission of the term $m = n = 0$ from the sum. As defined $f(z)$ is analytic for all Im $z > 0$ and has poles at every rational number on the real axis: $z = n/m$. Since any neighbourhood of any real number contains a rational number (*i.e.* the rational numbers are *dense* in the reals) it is impossible to define a finite-width strip that crosses the real axis and on which $f(z)$ is analytic. As a result the uniqueness theorem fails and there is no unique way to analytically extend $f(z)$ below the real axis.

3. Consider another such *Lacunary* function:

$$f(z) = z + z^2 + z^4 + z^8 + ... = \sum_{n=0}^{\infty} z^{2^n}. \tag{6.3.4}$$

$f(z)$ defined by eq.(6.3.4) for $|z| < 1$ diverges as $z \to 1$. Any analytic continuation of $f(z)$ therefore has a singular point at $z = 1$. Since $f(z)$ satisfies $f(z) = z + f(z^2)$ this also implies that $f(z)$ diverges for $z = -1$. Similarly, $f(z) = z + z^2 + f(z^4)$, implies that $z = \pm i$ are also singular. Repeating this argument shows that $f(z)$ diverges for every z on the unit circle that satisfies $z^{2^n} = 1$ for some n. Since these roots of unity are dense on the unit circle, there is no unique analytic extension of eq. (6.3.4) out of the unit disc.

6.4 Methods of Analytic Continuation

The previous theorem establishes the uniqueness of but does not address the construction of an analytic continuation of a given function. We present three methods of varying utility:

1. Continuation by Power Series:

   This technique is the method that is used in the proof of the uniqueness theorem. It is more useful in principle than in practice. Suppose a function is defined by its Taylor series about $z = z_0$ for which the radius of convergence is c:

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n, \qquad \text{for all } |z - z_0| < c \tag{6.4.1}$$

   In principle, the Taylor series gives the values of $f(z)$ and all of its derivatives at any point, z_1, within the disc of convergence:

$$\frac{d^k f}{dz^k}(z_1) = \sum_{n=0}^{\infty} a_n n(n-1)...(n-k+1)(z_1 - z_0)^{n-k}. \tag{6.4.2}$$

   These derivatives can therefore be used to define the Taylor series of $f(z)$ about $z = z_1$. The key point is that the series about $z = z_1$ then converges for a disc that need not be contained within the original domain of convergence of the series about $z = z_0$. In this way the definition of the function has been extended to the union of the discs of convergence of the two Taylor expansions.

   As an example, the function $f(z) = (1-z)^{-1}$ has the Taylor expansion: $f(z) = \sum_{n=0}^{\infty} z^n$ about $z = 0$. It converges for all $|z| < 1$. If the series is used to define the power series about $z = -\frac{1}{2}$ lying within this disc of convergence then the radius of convergence of the new series is $c = \frac{3}{2}$. The series definition would thereby be extended to the union of the two discs: $D_1 = \{z \ni |z| < 1\}$ and $D_2 = \{z \ni |z + \frac{1}{2}| < \frac{3}{2}\}$.

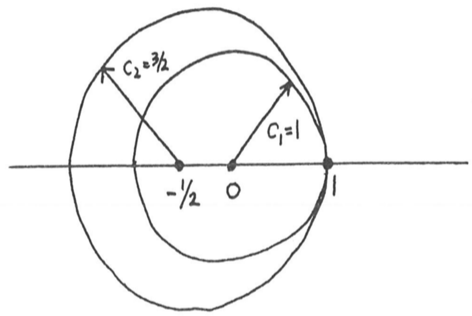

Figure 6.4. The Discs of Convergence of the Two Series

2. Schwarz' Reflection Principle

   If the domain in which the function is known intersects the real axis then its extension into the domain found by reflection through this axis is given

   by the following:

   Theorem 6.7. *Suppose $f(z)$ is a known analytic function on a connected open region, D, that has a single open line segment, L, of the real axis as a part of its boundary. Then, if $f(z)$ is real and continuous on L, and if D^* denotes the reflection of D about the real axis, then the continuation of $f(z)$ to $D \cup D^*$ is given by:*

$$f(z) = \begin{cases} f(z), & \text{for} \quad z \in D; \\ [f(z^*)]^*, & \text{for} \quad z \in D^*\,. \end{cases} \tag{6.4.3}$$

   Proof. We must show that $f(z)$ as defined in eq.(6.3.5) is analytic in D^* and across L. Analyticity in D^* follows immediately from that in D. Continuity across L follows because $f(z)$ was assumed to be real and continuous on L. Analyticity across L can then be shown using Morera's theorem, stated as point 13 of section (6.2) above. To see this break any loop within $D \cup D^*$ into a loop within D, including a segment along L, together with a loop in D^* that also includes the same segment along L but traversed in the opposite direction. The contour integral over any such loop is the sum of integrals over the loops in D and D^*. These vanish by Cauchy's theorem so the integral along any loop in $D \cup D^*$ also vanishes. Morera's theorem then implies the analyticity of $f(z)$. □

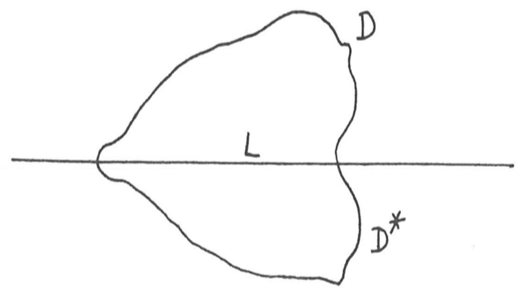

Figure 6.5. The Regions D and D^*

3. Continuation by Integration by Parts:

   This method is useful for functions that are defined by an integral representation, and is best illustrated by an example. Consider the function $\Gamma(z)$ defined in eq.(6.3.1). We rewrite the integral by integrating by parts:

$$\begin{aligned}\Gamma(z) &= \int_0^\infty t^{z-1} e^{-t}\, dt \\ &= \frac{1}{z}\int_0^\infty \frac{d}{dt}\left(t^z\right) e^{-t}\, dt \\ &= \frac{1}{z}\left[t^z e^{-t}\right]_0^\infty + \frac{1}{z}\int_0^\infty t^z e^{-t}\, dt.\end{aligned} \tag{6.4.4}$$

   The endpoint terms vanish in the region of definition, Re $z > 0$, and so can be neglected. The remaining integral is a representation for $\Gamma(z)$ in this region. This last representation has a larger domain of convergence, however, since the integral exists for all Re $z > -1$, and so it can be used as a definition of $\Gamma(z)$ for the strip $0 >$ Re $z > -1$. Because of the theorem guaranteeing the uniqueness of any analytic continuation, this new definition is the unique extension of $\Gamma(z)$ into this strip. Repeated integrations by parts can similarly be used to continue $\Gamma(z)$ to the rest of the complex plane.

6.5 Euler's Functions Gamma and Beta

The function $\Gamma(z)$, defined for Re $z > 0$ in eq.(6.3.1), was introduced by Euler and bears his name. Since it is ubiquitous in the solutions of the O.D.E.'s that we shall find, it is useful to build up some of its properties for future reference.

In the previous section we saw that $\Gamma(z)$ could be analytically continued away from its original domain of definition by successive integrations by parts. As seen from eq.(6.4.4), the result has the property that:

$$\Gamma(z) = \frac{\Gamma(z+1)}{z} = ... = \frac{\Gamma(z+n+1)}{z(z+1)\dots(z+n)} \tag{6.5.1}$$

for any positive integer, n. Since, by explicit evaluation, $\Gamma(1) = \int_0^\infty \exp(-t)\, dt = 1$, we see that property (6.5.1) ensures that $\Gamma(k+1) = k!$ for any positive integer, k. Similarly, since it was by definition analytic in the region Re $z > 0$, eq.(6.5.1) implies that $\Gamma(z)$ is analytic everywhere except for simple poles at $z = 0, -1, -2, ...$ Its residue at $z = -n$ is given by:

$$B_n = \lim_{z\to -n}[(z+n)\Gamma(z)] = \frac{\Gamma(1)}{(-n)(1-n)(2-n)\dots(-1)} = \frac{(-)^n}{n!}. \tag{6.5.2}$$

Another useful function of Euler's is the Beta function, $B(\alpha,\beta)$, defined by:

$$B(\alpha,\beta) = \int_0^1 (1-t)^{\alpha-1} t^{\beta-1}\, dt. \tag{6.5.3}$$

This definition converges provided that Re $\alpha > 0$ and Re $\beta > 0$. An alternative way to write eq.(6.5.3) is found by performing the change of variables $t = \sin^2\theta$ in terms of which:

$$B(\alpha,\beta) = 2\int_0^{\pi/2} \cos^{2\alpha-1}\theta \sin^{2\beta-1}\theta\, d\theta. \tag{6.5.4}$$

The next step is to extend this definition to the remainder of the complex plane. Rather than working directly with the integral representation we make a connection to $\Gamma(z)$. To do this consider the product:

$$\Gamma(\alpha)\Gamma(\beta) = \int_0^\infty du \int_0^\infty dv\; u^{\alpha-1}e^{-u}v^{\beta-1}e^{-v} \tag{6.5.5}$$

Changing variables to $u = x^2$ and $v = y^2$ and then switching to polar coordinates $x = r\cos\theta$ and $y = r\sin\theta$ gives the following:

$$\Gamma(\alpha)\Gamma(\beta) = 4\left[\int_0^\infty dr\; r^{2\alpha+2\beta-1}e^{-r^2}\right]\left[\int_0^{\pi/2} d\theta\;\cos^{2\alpha-1}\theta\sin^{2\beta-1}\theta\right] \tag{6.5.6}$$

Changing variables to $t = r^2$ and comparing to the definition (6.3.1) shows that the first factor is equal to $\frac{1}{2}\Gamma(\alpha+\beta)$. The second factor, on the other hand, when compared to eq.(6.5.4), is equal to $\frac{1}{2}B(\alpha,\beta)$. Combining these relations and solving for $B(\alpha,\beta)$ gives the equation:

$$B(\alpha,\beta) = B(\beta,\alpha) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \tag{6.5.7}$$

Although this last equation has been derived for positive Re α and Re β, for which the integral representations converge, it must hold for all complex α and β. This is because both sides of eq.(6.5.7) are analytic and the left-hand-side agrees with the definition, (6.5.2) within its domain of validity. Since the analytic continuation of eq.(6.5.2) is unique, it must be given by eq.(6.5.7) away from the original domain of definition.

In order to further illustrate complex-variable techniques and for further use we now develop some properties of the Γ-function.

1. Reflection Formula:

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)} \tag{6.5.8}$$

In order to prove the validity of eq.(6.5.8) consider the beta-function evaluated at $\alpha = z$ and $\beta = 1 - z$. Using $\Gamma(1) = 1$ gives:

$$\Gamma(z)\Gamma(1-z) = B(1-z, z) = \int_0^1 t^{z-1}(1-t)^{-z}\, dt \tag{6.5.9}$$

in which the final integral representation converges for $0 < \mathrm{Re}\ z < 1$. The logic is to use this integral representation to prove relation (6.5.8) within the strip in which it converges and then to use the uniqueness of analytic continuation to infer that the result also holds in the remainder of the complex plane. The change of variables $t = u/(u+1)$ in (6.5.9) produces the following integral:

$$B(z, 1-z) = \int_0^\infty \frac{u^{z-1}}{1+u}\, du. \tag{6.5.10}$$

The idea is to explicitly evaluate this integral by residues. In order to do so we need to know the singularities of the integrand. By inspection the integrand has a simple pole at $u = -1$ and, for z not an integer, a branch point at $u = 0$. We can choose to define $u^z = \exp(z \log u)$ with $\log u = \log |u| + i \arg u$ and $0 < \arg u < 2\pi$. This amounts to choosing the branch cut to lie along the positive real axis.

Consider, then, the integral (6.5.10) evaluated along the contour $C = \{z \ni z = -1 + \frac{1}{2}e^{2\pi i t}\}$ consisting of a circle with radius $\frac{1}{2}$ centered on the pole at $u = -1$ and traversed in the counterclockwise sense. This integral is easily evaluated by residues and gives:

$$I \equiv \int_C \frac{u^{z-1}}{1+u}\, du = -2\pi i \exp(i\pi z). \tag{6.5.11}$$

The argument now relies on deforming the contour, C, to a new contour, C', that is related to the range of integration of the integral in eq.(6.5.10). The new contour is given by $C' = C_\infty \cup C_- \cup C_0 \cup C_+$ where the segments are defined as the limit of the following contours as $R \to \infty$ and $r \to 0$:

$$\begin{aligned} C_\infty &= \{u = Re^{2\pi i t},\ 0 \le t \le 1\}; \\ C_- &= \{u = R(1-t) - ir,\ 0 \le t \le 1\}; \\ C_0 &= \{u = re^{i\pi(\frac{3}{2}-t)},\ 0 \le t \le 1\}; \\ C_+ &= \{u = Rt + ir,\ 0 \le t \le 1\}; \end{aligned} \tag{6.5.12}$$

The claim is that the contribution to the integral from the contours C_∞ and C_0 vanish in the limit that $r \to 0$ and $R \to \infty$. The sole contribution to the integral along C' is

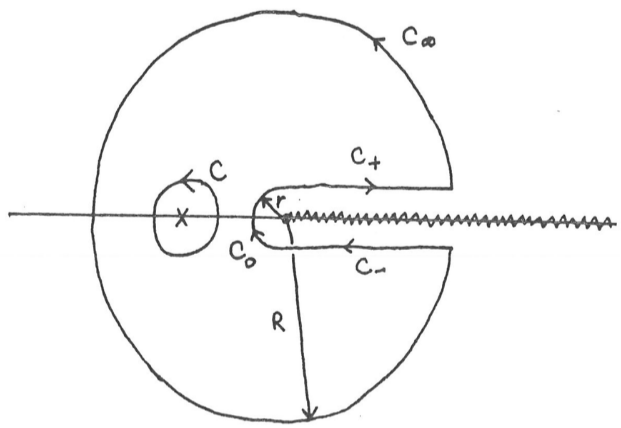

Figure 6.6. The Contours Referred to in the Proof

therefore from the integrals on either side of the branch cut, $C_{\pm}$. These integrals are related to the integral (6.5.10) that we wish to compute.

The first step is to show that the arcs 'at infinity' and 'at zero' contribute nothing. Consider the contribution of C_∞. We have:

$$\begin{aligned} A &\equiv \lim_{R\to\infty} A(R) \\ \text{with} \qquad A(R) &= \int_0^1 \frac{R^{z-1}e^{2\pi i t(z-1)}}{1+Re^{2\pi i t}}\, Re^{2\pi i t} 2\pi i\, dt \\ &= 2\pi i R^z \int_0^1 \frac{e^{2\pi i t z}}{1+Re^{2\pi i t}}\, dt \end{aligned} \tag{6.5.13}$$

The vanishing of this limit is proven by bounding the absolute value $|A(R)|$ from above by a function that vanishes as $R \to \infty$. The desired inequality follows from repeated use of the triangle inequality.

$$\begin{aligned} |A(R)| &\leq 2\pi R^{\text{Re } z} \int_0^1 \frac{e^{-2\pi t \text{Im } z}}{|1+Re^{2\pi i t}|}\, dt \\ &\leq \frac{2\pi R^{\text{Re } z}}{R-1} \int_0^1 e^{-2\pi t \text{Im } z}\, dt \end{aligned} \tag{6.5.14}$$

The first inequality is an application of the triangle inequality in its integral form: $\left|\int f(t)\, dt\right| \leq \int |f(t)|\, dt$. The second follows from the triangle inequality applied to the denominator of the integrand: $|1 + R\exp(2\pi i t)| \geq R - 1$. The remaining integral in the final line of eq. (6.5.14) is elementary. It remains to perform the limit $R \to \infty$. The key observation is that we have required that z lie in the strip $0 < \text{Re } z < 1$. With this choice the right-hand-side of eq. (6.5.14) vanishes when R becomes arbitrarily large, so $A = \lim_{R\to\infty} A(R) = 0$ as required.

An identical argument may be applied to the semicircular arc near the origin:

$$
\begin{aligned}
B &= \lim_{r\to 0} B(r) \\
\text{with} \qquad B(r) &= -i\pi r^z \int_0^1 \frac{e^{i\pi z\left(\frac{3}{2}-t\right)}}{1+re^{i\pi\left(\frac{3}{2}-t\right)}}\, dt \\
\text{so} \qquad |B(r)| &\le \frac{\pi r^{\mathrm{Re}\ z}}{1-r} \int_0^1 e^{-\pi \mathrm{Im}\ z\left(\frac{3}{2}-t\right)}\, dt
\end{aligned} \tag{6.5.15}
$$

The limit as $r \to 0$ of the right-hand-side of the last equation in (6.5.15) is zero for any $0 < \mathrm{Re}\ z < 1$. This demonstrates that $B(r)$ vanishes in this limit and so, therefore, does B.

The remaining contribution to the contour integral around C' is from the segments just above and below the positive real axis. This gives:

$$
\begin{aligned}
I &= \int_{C'} \frac{u^{z-1}}{1+u}\, du \\
&= \int_{C_+} \frac{u^{z-1}}{1+u}\, du + \int_{C_-} \frac{u^{z-1}}{1+u}\, du \\
&= \lim_{r\to 0} \int_0^\infty \left[\frac{(t+ir)^{z-1}}{1+t+ir} - \frac{(t-ir)^{z-1}}{1+t-ir} \right] dt \\
&= \left[1 - e^{2\pi i z}\right] \int_0^\infty \frac{t^{z-1}}{1+t}\, dt
\end{aligned} \tag{6.5.16}
$$

Notice that the integral on the right-hand-side of this final equation is precisely the integral of eq. (6.5.10) that we wish to evaluate. Combining, then, eqs. (6.5.10), (6.5.11) and (6.5.16) allows the desired conclusion:

$$
\begin{aligned}
\Gamma(z)\Gamma(1-z) &= \int_0^\infty \frac{t^{z-1}}{1+t}\, dt \\
&= \frac{1}{1-e^{2\pi i z}} \int_{C'} \frac{u^{z-1}}{1+u}\, du \\
&= \frac{-2\pi i e^{i\pi z}}{1-e^{2\pi i z}} \\
&= \frac{\pi}{\sin(\pi z)}.
\end{aligned} \tag{6.5.17}
$$

As indicated earlier, even though this result has only been proven for z lying within the strip $0 < \mathrm{Re}\ z < 1$ its validity follows for the rest of the complex z-plane by virtue of the uniqueness of analytic continuation. That is, both sides of eq. (6.5.17) are analytic apart from isolated poles throughout the plane and they agree within the strip $0 < \mathrm{Re}\ z < 1$. They must therefore agree throughout the rest of the plane.

An important special case of the reflection formula is obtained by choosing $z = \frac{1}{2}$. In this case it reads $[\Gamma(\frac{1}{2})]^2 = \pi$. Since the integral formula (6.4.4) implies that $\Gamma(\frac{1}{2}) > 0$, it follows that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Repeated use of property $\Gamma(z+1) = z\Gamma(z)$ that follows from eq. (6.5.1) gives, for Γ evaluated at any half-odd number [1]:

$$\Gamma\left(n+\frac{1}{2}\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}. \tag{6.5.18}$$

The next property to be proven for the Γ-function relates its value at $2z$ to its value at z and $z+\frac{1}{2}$.

2. Duplication Formula:

$$\Gamma(z)\Gamma\left(z+\frac{1}{2}\right) = 2^{1-2z}\sqrt{\pi}\,\Gamma(2z) \tag{6.5.19}$$

This result follows from an artful choice of variables in the integral representation of the beta-function. Consider, for Re $z > 0$:

$$B(z,z) = \frac{[\Gamma(z)]^2}{\Gamma(2z)} = \int_0^1 t^{z-1}(1-t)^{z-1}\,dt. \tag{6.5.20}$$

Change variables to $v = (2t-1)^2$. As t varies from 0 to 1 v covers the region from 0 to 1 twice. The resulting integral is:

$$\begin{aligned} B(z,z) &= 2^{1-2z}\int_0^1 v^{-\frac{1}{2}}(1-v)^{z-1}\,dv \\ &= 2^{1-2z}B\left(\frac{1}{2},z\right) = 2^{1-2z}\frac{\Gamma(z)\Gamma(\frac{1}{2})}{\Gamma(z+\frac{1}{2})}. \end{aligned} \tag{6.5.21}$$

The duplication formula is now obtained by multiplying eq. (6.5.21) through by $\Gamma(2z)\Gamma(z+\frac{1}{2})/\Gamma(z)$. As usual, although the proof of this result used the assumption that Re $z > 0$, both sides of the final equation are analytic and so the uniqueness theorem for analytic continuation ensures that the duplication formula holds for all z.

The final property of Euler's Γ-function that is to be proven is *Stirling's Formula* and is the subject of the next chapter.

[1] $(2n-1)!! = (2n-1)(2n-3)\cdots 3\cdot 1$.

7 Asymptotic Forms and the Method of Steepest Descent

For many purposes it is useful to be able to describe the behaviour of a given function for large values of its argument. It is of particular use to be able to do so even when the function is defined by a series or integral representation, as are the solutions we construct to the O.D.E.'s in later chapters. For integral representations a very useful technique for generating such *asymptotic forms* is the method of *Steepest Descent*. (This method is also often referred to as the *Stationary-Phase Approximation* or the *Saddle-Point Approximation*.) This technique is described in this chapter in its simplest possible setting. It is worth keeping in mind, however, that the technique has a more general setting and is much more powerful than might appear from the presentation below.

7.1 The Approximation

We encounter the saddle-point approximation within the context of deriving the behaviour of $\Gamma(z)$ for large values of z. The result is expressed by
Stirling's Formula:

$$\Gamma(z+1) = \sqrt{2\pi} z^{z+\frac{1}{2}} e^{-z} \left[1 + \frac{1}{12z} + \frac{1}{288z^2} + O(z^{-3})\right] \tag{7.1.1}$$

Formula (7.1.1) is derived with the help of the integral representation, (6.3.1), that defines $\Gamma(z)$ and is applicable for Re $z > 0$. The first step in doing so is to change variables to $u = t/z$ and rotate the resulting contour of integration back to the positive real axis.

$$\begin{aligned} \Gamma(z+1) &= \int_0^\infty t^z e^{-t}\, dt \\ &= z^{z+1} \int_0^\infty e^{z(\log u - u)}\, du \end{aligned} \tag{7.1.2}$$

This form for the integrand is the one for which the method of steepest descent may apply, namely to integrals of the form:

$$I(z) = \int_0^\infty e^{zf(u)}\, du. \tag{7.1.3}$$

In the case at hand the real function $f(u)$ is given by $f(u) = \log u - u$. In this more general case we wish to know the behaviour of $I(z)$ as $z \to \infty$. (In general it is not necessary, for the method of steepest descent, for $f(u)$ to be real or for the integration range to consist only of the positive real axis.) $f(u)$ must go to $-\infty$ sufficiently quickly as $u \to \infty$ if this integral to converge. We also assume that $f(u)$ also has a single stationary point which is a global maximum for some $u \neq 0$. This last assumption is again not necessary but is chosen to simplify the argument.

It is easy to see that this function has a single stationary point at $u = 1$ where $f'(1) = [u^{-1} - 1]|_{u=1} = 0$ and $f(1) = -1$. This corresponds to a global maximum. The function falls off to negative infinity as $u \to 0$ and as $u \to \infty$, and the second derivative is $f''(u) = -1/u^2 < 0$ so it is everywhere concave downwards. The Taylor expansion of $f(u)$ about $u = 1$ is elementary and is given for $|u - 1| < 1$ by:

$$f(u) = -1 - \sum_{n=2}^{\infty} \frac{(-)^n}{n} (u-1)^n. \tag{7.1.4}$$

Here's the essence of the argument. For larger and larger Re z the integrand, $e^{zf(u)}$, is more and more strongly peaked about the stationary point of $f(u)$ at $u = 1$. It should therefore follow that the behaviour of $I(z)$ for such large z is completely governed by the behaviour of the integrand in the neighbourhood of the stationary point $u = 1$. A good approximation should therefore be obtained by Taylor-expanding the argument of the exponent about this point, using eq. (7.1.4). The first term of this series that depends on $u - 1$ is quadratic. The idea is to leave this term in the exponent to ensure that the integral over u converges for large u but to again Taylor-expand the exponential of the remainder of the series and to then integrate the result term-by-term. The final claim is that the resulting function of z so obtained furnishes an asymptotic series for $I(z)$ of the form:

$$I(z) = A(z) \left[1 + \frac{B}{z} + O(z^{-2}) \right]. \tag{7.1.5}$$

We illustrate this method by using it to calculate Stirling's formula, (7.1.1). A demonstration that the terms neglected in the derivation are negligible, *i.e.* that the result is an asymptotic series, is sketched at the end.

Writing the Taylor expansion of eq. (7.1.4) as $f(u) = -1 - \frac{1}{2}(u-1)^2 - S(u)$, in which $S(u) = \sum_{n=3}^{\infty} \frac{(-)^n}{n} (u-1)^n$ denotes all of the terms beyond second order in $(u-1)$, allows the argument of the integrand to be written:

$$\begin{aligned}
\exp[zf(u)] &= \exp\left[-z - \frac{1}{2} z (u-1)^2 - zS(u) \right] \\
&= e^{-z} \exp\left[-\frac{1}{2} z (u-1)^2 \right] \exp\left[-zS(u) \right] \\
&= e^{-z} \exp\left[-\frac{1}{2} z (u-1)^2 \right] \sum_{k=0}^{\infty} \frac{1}{k!} [-zS(u)]^k \\
&= e^{-z} \exp\left[-\frac{1}{2} z (u-1)^2 \right] \left[1 - z \left(\sum_{n=3}^{\infty} \frac{(-)^n}{n} (u-1)^n \right) \right. \\
&\qquad \left. + \frac{1}{2} z^2 \left(\sum_{n=3}^{\infty} \frac{(-)^n}{n} (u-1)^n \right)^2 + \ldots \right]
\end{aligned} \tag{7.1.6}$$

$I(z)$ is to be constructed by integrating the last result of eqs. (7.1.6) term-by-term. This gives the unwieldy result:

$$I(z) = e^{-z}\left[\int_0^\infty e^{-\frac{1}{2}z(u-1)^2}\,du - z\sum_{n=3}^{\infty}\frac{(-)^n}{n}\int_0^\infty e^{-\frac{1}{2}z(u-1)^2}(u-1)^n\,du + \right.$$
$$\left. +\frac{1}{2}z^2\sum_{n=3}^{\infty}\sum_{m=3}^{\infty}\frac{(-)^{n+m}}{mn}\int_0^\infty e^{-\frac{1}{2}z(u-1)^2}(u-1)^{n+m}\,du + ...\right] \tag{7.1.7}$$

Since the integral is dominated, for large Re z, by the behaviour of the integrand near $u = 1$ the lower limit of integration may be taken to be negative infinity rather than zero with negligible error. (More about the size of this error later.) The integrals are then straightforward and may be evaluated in terms of Euler's Γ-function. Using the variable $t = \frac{1}{2}z(u-1)^2$ and rotating the contour back onto the real axis gives:

$$\begin{aligned}\int_{-\infty}^\infty e^{-\frac{1}{2}z(u-1)^2}(u-1)^p\,du &= 0 \qquad \text{if} \qquad p = 2n+1\\ &= \left(\frac{2}{z}\right)^{n+\frac{1}{2}}\int_0^\infty e^{-t}t^{n-\frac{1}{2}}\,dt \qquad \text{if} \qquad p = 2n\\ &= \left(\frac{2}{z}\right)^{n+\frac{1}{2}}\Gamma\left(n+\frac{1}{2}\right)\end{aligned} \tag{7.1.8}$$

Using eq. (7.1.8) in eq. (7.1.7) gives:

$$I(z) = \left(\frac{2}{z}\right)^{\frac{1}{2}}e^{-z}\left\{\Gamma(\frac{1}{2}) - z\sum_{n=2}^{\infty}\left(\frac{1}{2n}\right)\left(\frac{2}{z}\right)^n\Gamma\left(n+\frac{1}{2}\right) + \right.$$
$$\left. +\frac{1}{2}z^2\sum_{n=3}^{\infty}{\sum_{m=3}^{\infty}}'\left(\frac{1}{mn}\right)\left(\frac{z}{2}\right)^{-\frac{1}{2}(m+n)}\Gamma\left[\frac{1}{2}(m+n+1)\right] + ...\right\} \tag{7.1.9}$$

The prime on the sum indicates that only those terms with $n+m$ even are to be included.

Collecting the first few powers of $1/z$ and using eq. (6.5.18) for $\Gamma(p+\frac{1}{2})$ gives:

$$\begin{aligned}I(z) &= \left(\frac{2\pi}{z}\right)^{\frac{1}{2}}e^{-z}\left[1 - \sum_{n=2}^{\infty}\frac{(2n-1)!!}{2n}z^{1-n} + \right.\\ &\qquad \left. + \sum_{n=3}^{\infty}{\sum_{m=3}^{\infty}}'\frac{(m+n-1)!!}{2mn}z^{2-\frac{1}{2}(m+n)} + ...\right]\\ &= \left(\frac{2\pi}{z}\right)^{\frac{1}{2}}e^{-z}\left[1 + \frac{1}{12z} + \frac{1}{288z^2} + ...\right]\end{aligned} \tag{7.1.10}$$

Inserting this asymptotic form for $I(z)$ into $\Gamma(z+1) = z^{z+1}I(z)$ reproduces the first terms of Stirling's formula.

What remains to be shown is that the approximations employed in deriving eq. (7.1.10) only contribute terms that are small. The claim is that the inclusion of higher powers of $(u-1)$ in the series for $\exp[-zS(u)]$ only adds corrections to (7.1.10) that are $O(z^{-3})$. The remaining approximations, such as that of changing the lower limit of integration in eq. (7.1.7) to negative infinity, introduce an error that is smaller than any power of z: *i.e.* they are $o(z^{-n})$ for any n.

7.2 The Accuracy of the Approximation

The claim to be demonstrated is that the series (7.1.10) in $1/z$ is an asymptotic series for the function $I(z)$. That is, it remains to prove the following:

$$\begin{aligned} A_1(z) &\equiv \frac{I(z)}{\sqrt{\frac{2\pi}{z}}e^{-z}} - 1 = O(z^{-1}) \\ A_2(z) &\equiv A_1(z) - \frac{1}{12z} = O(z^{-2}) \end{aligned} \tag{7.2.1}$$

and so on.

Proof. In order to simplify the argument and to avoid a mathematical technicality a slightly easier proposition than this will be argued. The simplification is to establish eq. (7.2.1a), say, for the proper integral

$$A_1[0,R;z] \equiv \sqrt{\frac{z}{2\pi}}e^{-zf(u_0)} \int_0^R e^{zf(u)}\,du \tag{7.2.2}$$

rather than the original improper integral in which the upper limit, R, is taken to infinity. The argument in the more general case of interest requires some modification but is basically equivalent.

The result is demonstrated by dividing the integration into two regions: one lying within a neighbourhood $(u_0-\epsilon, u_0+\epsilon)$ surrounding the stationary point u_0 and the other consisting of the remainder of the integration region:

$$A_1[0,R;z] = A_1[0,u_0-\epsilon;z] + A_1[u_0-\epsilon,u_0+\epsilon;z] + A_1[u_0+\epsilon,R;z]. \tag{7.2.3}$$

ϵ is chosen smaller than the radius of convergence of the Taylor series of $f(u)$ about the point u_0. This allows the integrand of the second term on the right-hand-side to be expanded in powers of $(u-u_0)$. There are then four steps in the argument.

1. The contribution of the first and the last terms on the right-hand-side are shown to be exponentially small as $z \to \infty$. More precisely they are shown to be $o(z^{-p})$ for any positive p.

2. Next, the integrand of the remaining integral from $u_0-\epsilon$ to $u_0+\epsilon$ is Taylor expanded about the stationary point, u_0.

3. Each term in the term-by-term integration of this series has a *gaussian* form, being the integral of the product of an exponential of a square, $(u-u_0)^2$, multiplied by a power of $(u-u_0)$. The argument proceeds by showing that the effect of changing the limits of these integrals to $-\infty$ and $+\infty$ is to neglect terms which are also exponentially small.

4. Finally, the remaining gaussian integrals from $-\infty$ to $+\infty$ are explicitly performed and furnish the asymptotic series (7.1.10).

To establish the first part, consider:

$$A_1[0, u_0-\epsilon; z] = \sqrt{\frac{z}{2\pi}} e^{-zf(u_0)} \int_0^{u_0-\epsilon} e^{zf(u)}\, du. \tag{7.2.4}$$

An inequality is easily established for $|I[0, u_0-\epsilon; z]|$ using the triangle inequality discussed following eq. (6.5.14). That is:

$$\begin{aligned} |A_1[0, u_0-\epsilon; z]| &\le \sqrt{\frac{|z|}{2\pi}} e^{-f(u_0)\mathrm{Re}\, z} \int_0^{u_0-\epsilon} e^{f(u)\mathrm{Re}\, z}\, du \\ &\le \sqrt{\frac{|z|}{2\pi}} e^{-f(u_0)\mathrm{Re}\, z} \int_0^{u_0-\epsilon} e^{f(u_0-\epsilon)\mathrm{Re}\, z}\, du \\ &= \sqrt{\frac{|z|}{2\pi}} (u_0-\epsilon) e^{-[f(u_0)-f(u_0-\epsilon)]\mathrm{Re}\, z}. \end{aligned} \tag{7.2.5}$$

In the second inequality the property that $f(u)$ has a global maximum at u_0 has been used to imply $f(u) \le f(u_0-\epsilon)$ throughout the region, $(0, u_0-\epsilon)$, of integration. The argument is completed with the realization that, by assumption, Re $z > 0$ and that $f(u_0) > f(u_0-\epsilon)$ so the coefficient of Re z in the exponent of the last line is negative. It follows therefore that

$$\lim_{\mathrm{Re}\, z\to\infty} [z^p A_1[0, u_0-\epsilon; z]] = 0 \tag{7.2.6}$$

for any positive p and ϵ so $A_1[0, u_0-\epsilon; z] = o(z^{-p})$. An identical argument also shows that

$$|A_1[u_0+\epsilon, R; z]| \le \sqrt{\frac{|z|}{2\pi}} (R-u_0-\epsilon) e^{-[f(u_0)-f(u_0-\epsilon)]\mathrm{Re}\, z} \tag{7.2.7}$$

and so, for any fixed R, $A_1[u_0+\epsilon, R; z] = o(z^{-p})$ for any positive p. The contribution of the segments $(0, u_0-\epsilon)$ and $(u_0+\epsilon, R)$ to A_1 is therefore exponentially small for large Re z.

The dominant part of the integral therefore comes from the immediate neighbourhood of the stationary point, u_0. The next step is to rewrite this contribution as a series of gaussian integrals. To do so Taylor expand $f(u)$:

$$f(u) = \sum_{n=0}^{\infty} a_n v^n \tag{7.2.8}$$

where $v = u - u_0$ defines the variable v. The result for A_1 is:

$$
\begin{aligned}
A_1[u_0 - \epsilon, u_0 + \epsilon; z] &= \sqrt{\frac{z}{2\pi}} e^{-f(u_0)z} \int_{-\epsilon}^{\epsilon} \exp\left[z \left(f(u_0) + \sum_{n=2}^{\infty} a_n v^n \right) \right] \, dv \\
&= \sqrt{\frac{z}{2\pi}} \int_{-\epsilon}^{\epsilon} e^{\frac{1}{2} z v^2 f''(u_0)} \left[1 + \sum_{n=3}^{\infty} b_n v^n \right]. \qquad (7.2.9)
\end{aligned}
$$

where the coefficients b_n are given in terms of the Taylor-series a_n by

$$
\begin{aligned}
b_3 &= z a_3 \\
b_4 &= z a_4 \\
b_5 &= z a_5 \\
b_6 &= z a_6 + \frac{1}{2} z^2 a_3^2 \\
b_7 &= z a_7 + z^2 a_3 a_4 \qquad (7.2.10)
\end{aligned}
$$

etc.

The first observation is that all terms in the last of eq. (7.2.9) involving an odd power of v vanish because the integration region is symmetric about $v = 0$.

The next point is that the limits of integration in the remaining terms can be changed from $\pm\epsilon$ to $\pm\infty$ up to a correction that vanishes exponentially quickly as $\mathrm{Re}\, z \to \infty$. This is established in the same way as was done for the contribution of the interval $0 \le u \le u_0 - \epsilon$. That is, the following limit:

$$
\lim_{\mathrm{Re}\, z \to \infty} \left[z^p \int_{-\infty}^{-\epsilon} e^{\frac{1}{2} z v^2 f''(u_0)} v^n \, dv \right] = 0 \qquad (7.2.11)
$$

may be proven for any positive p and ϵ (exercise!). Ditto for the contribution from (ϵ, ∞).

We are left with the result:

$$
\begin{aligned}
A_1(z) &= \sqrt{\frac{z}{2\pi}} e^{-f(u_0)z} \int_0^R e^{z f(u)} \, du \\
&= \sqrt{\frac{z}{2\pi}} \int_{-\infty}^{\infty} e^{\frac{1}{2} z v^2 f''(u_0)} \left[1 + \sum_{n=2}^{\infty} b_{2n} v^{2n} \right] + o(z^{-p})
\end{aligned}
\qquad (7.2.12)
$$

for any positive p. The z-dependence of this series is seen by changing variables to $w = v\sqrt{z}$. Eq. (7.2.12) becomes:

$$
A_1(z) = \left[1 + \frac{1}{\sqrt{2\pi}} \sum_{n=2}^{\infty} \frac{b_{2n}}{z^n} \int_{-\infty}^{\infty} e^{\frac{1}{2} w^2 f''(u_0)} w^{2n} \right] + o(z^{-p}). \qquad (7.2.13)
$$

Notice that b_{2n} is a polynomial of z of degree $< n$. Therefore the combination b_{2n}/z^n is at least $O(z^{-1})$. Evaluation of this series explicitly reproduces the desired result (7.1.10). This completes the proof. □

Corrolary 7.1. *There is a simple generalization of the steepest-descent results to the case where the integrand has the form:*

$$I(z) = \int e^{zf(u)} g(u)\, du. \tag{7.2.14}$$

The asymptotic form in this case is given as before by expanding the integrand about the stationary point, u_0, of the exponent:

$$I(z) = \sqrt{\frac{2\pi}{zf''(u_0)}} e^{zf(u_0)} g(u_0)[1 + O(z^{-1})]. \tag{7.2.15}$$

The proof of this result proceeds exactly as above.

8 Power-Series Solutions

8.1 Introduction

With the method of separation of variables the problem of solving for solutions to P.D.E.'s with the separated form has reduced to solving a system of linear second-order O.D.E.'s. The generic form for such an equation is:

$$y''(z) + p(z)y'(z) + q(z)y(z) = 0. \tag{8.1.1}$$

Examples of the types of equations so obtained are Bessel's equation and the Associated Legendre equation already encountered in chapter 5. These equations usually do not admit solutions in closed form in terms of elementary functions. The present chapter is devoted to a general method of solution that can be applied to all of the O.D.E.'s encountered by separation of variables.

The method is the technique of power-series solution. In this approach a power-series solution is assumed and its coefficients are determined by the requirement that the series satisfy the O.D.E.. Unfortunately it is not necessarily true that the solution to a given differential equation is analytic at any given point, so it need not be true that a solution exists of the form of an ordinary Taylor series. The key question then is to determine under what circumstances a given series will be guaranteed to furnish a solution to a given O.D.E. and to settle what form of series *ansatz* will always generate a solution.

Since these questions revolve around the analytic properties of the solutions, it is first necessary to derive some results that relate the behaviour of solutions to the analytic properties of the coefficient functions $p(z)$ and $q(z)$ of eq. (8.1.1).

8.2 Ordinary Points

The first useful property along these lines is the fact that the solutions to (8.1.1) are analytic at any point, z, at which both $p(z)$ and $q(z)$ are analytic. This is established by directly constructing a series form for the equation and seeing that it converges. A byproduct of this construction is the establishment of the existence of solutions, which was the missing step in the proof that the space of solutions to (8.1.1) is two-dimensional.

To start, a definition:

Definition 8.1. *Ordinary Point of an O.D.E.: A point, z_0, is said to be an ordinary point of the differential equation (8.1.1) if both $p(z)$ and $q(z)$ are analytic at $z = z_0$.*

Similarly,

Definition 8.2. *Singular Point of an O.D.E.: Any point, z_0, of an O.D.E. that is not ordinary is said to be a singular point of that O.D.E..*

The utility of these definitions lies in the following:

Theorem 8.1. *If z_0 is an ordinary point of eq. (8.1.1), then any of its solutions, $y(z)$, are analytic at $z = z_0$. They therefore admit uniformly convergent solutions of the form:*

$$y(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{8.2.1}$$

in some neighbourhood of z_0.

Proof. This result is established by rewriting the given second-order O.D.E. as a system of coupled first-order O.D.E.'s that may be formally solved. (Unfortunately the solution is not given in a form that is particularly useful for practical calculations.) To do so, perform the following change of dependent variable:

$$w(z) = \exp\left[-\frac{1}{2}\int_{z_0}^{z} p(u)\, du\right] y(z) \tag{8.2.2}$$

in terms of which the O.D.E. (8.1.1) becomes:

$$w'' + r(z) w = 0 \tag{8.2.3}$$

with:

$$r(z) = q(z) - \frac{1}{2} p'(z) - \frac{1}{4} p^2(z). \tag{8.2.4}$$

Define, now the second dependent variable by $v(z) = w'(z)$. Equation (8.1.1) is then equivalent to the system:

$$\begin{pmatrix} w'(z) \\ v'(z) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -r(z) & 0 \end{pmatrix} \begin{pmatrix} w(z) \\ v(z) \end{pmatrix} = \mathbf{A}(z) \begin{pmatrix} w(z) \\ v(z) \end{pmatrix} \tag{8.2.5}$$

in which the last equality defines the matrix function $\mathbf{A}(z)$.

The solution to this first-order problem is given by the useful strategem of converting the differential equation into an equivalent *integral equation*. To this end notice that the solution can be written in terms of initial conditions specified at the point z_0 by the following matrix relation:

$$\begin{pmatrix} w(z) \\ v(z) \end{pmatrix} = \mathbf{U}(z, z_0) \begin{pmatrix} w(z_0) \\ v(z_0) \end{pmatrix}. \tag{8.2.6}$$

Here $\mathbf{U}(z, z_0)$ is a matrix whose elements are functions of both the 'initial point', z_0, and the 'final point', z. Clearly $\mathbf{U}(z = z_0, z_0)$ is equal to the unit matrix, $\mathbf{I}$. Eq. (8.2.6) really expresses nothing beyond the fact that the solution to the system (8.2.5) depends linearly on the initial values specified at $z = z_0$. Plugging eq. (8.2.6) into the differential equation (8.2.5) gives a differential equation for the matrix $\mathbf{U}(z, z_0)$:

$$\frac{\partial \mathbf{U}}{\partial z}(z, z_0) = \mathbf{A}(z) \mathbf{U}(z, z_0). \tag{8.2.7}$$

Integrating eq. (8.2.7) with respect to z and using the initial conditions gives the desired integral equation for $\mathbf{U}$:

$$\mathbf{U}(z, z_0) = \mathbf{I} + \int_{z_0}^{z} \mathbf{A}(u)\mathbf{U}(u, z_0)\, du. \tag{8.2.8}$$

The advantage of rewriting the O.D.E. in this way is that it may now be solved by iteration. That is, substitute the left-hand-side of (8.2.8) back into the right-hand-side repeatedly, thereby generating an expansion for $\mathbf{U}(z, z_0)$ in powers of the matrix $\mathbf{A}(z)$:

$$\begin{aligned}
\mathbf{U}(z, z_0) &= \mathbf{I} + \int_{z_0}^{z} \mathbf{A}(u)\mathbf{U}(u, z_0)\, du \\
&= \mathbf{I} + \int_{z_0}^{z} \mathbf{A}(u)\, du + \int_{z_0}^{z} du \int_{z_0}^{u} dv\, \mathbf{A}(u)\mathbf{A}(v)\mathbf{U}(v, z_0) \\
&= \mathbf{I} + \sum_{n=1}^{\infty} \int_{z_0}^{z} du_1 \int_{z_0}^{u_1} du_2 \cdots \int_{z_0}^{u_{n-1}} du_n\, \mathbf{A}(u_1)\mathbf{A}(u_2)\cdots\mathbf{A}(u_n).
\end{aligned} \tag{8.2.9}$$

Choose now $z = z'$ where z' is an ordinary point of the O.D.E.. Then $r(u)$ is ensured to be analytic throughout some neighbourhood of z' and so will be analytic on the integration contour from z_0 to z' if z_0 is chosen sufficiently near to z'. The integrals in eqs. (8.2.9) are then guaranteed to converge and so each term of the series is a matrix whose entries are analytic functions near z'.

The issue to be established is whether this series converges in some neighbourhood of z'. The proof of this fact is given using the following series of lemmas.

Definition 8.3. *The key tool in this regard is the norm of a matrix, $\|\mathbf{M}\|$, given by:*

$$\|\mathbf{M}\|^2 \equiv |m_{11}|^2 + |m_{12}|^2 + |m_{21}|^2 + |m_{22}|^2 \tag{8.2.10}$$

for the matrix $\mathbf{M}$ defined by:

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}. \tag{8.2.11}$$

Lemma 8.1. *The utility of this definition is that a series of matrices: $\sum_n \mathbf{M}_n$ is guaranteed to converge if the sum of the norms: $\sum_n \|\mathbf{M}_n\|^2$ converges. A matrix sum is defined to converge if the sum defining each of its elements does. Now notice that the absolute value of any element of a matrix is less than or equal to the norm of the entire matrix. Therefore, if the norm of a sum of matrices converges, then so must the sum defining each of its elements.*

Lemma 8.2. *Next notice the following simple properties of the norm that can be established by direct application of the definitions:*

$$\begin{aligned}
\|\mathbf{M}\mathbf{N}\|^2 &\leq \|\mathbf{M}\|^2\|\mathbf{N}\|^2. \\
\|\mathbf{M} + \mathbf{N}\|^2 &\leq \|\mathbf{M}\|^2 + \|\mathbf{N}\|^2.
\end{aligned} \tag{8.2.12}$$

Here comes the main argument. We know that the elements of the matrix $\mathbf{A}(z)$ are analytic along the integration contour from z_0 to z'. This means that $\|\mathbf{A}(z)\|^2$ is bounded by some positive K for all z on the contour. It follows then that $\|\mathbf{A}(u_1)\cdots\mathbf{A}(u_n)\|^2$ is bounded by K^n. But then the norm of the series in eq. (8.2.9) is bounded from above by:

$$
\begin{aligned}
S &\equiv \| \sum_{n=0}^{\infty} \int_{z_0}^{z} du_1 \int_{z_0}^{u_1} du_2 \cdots \int_{z_0}^{u_{n-1}} du_n \; \mathbf{A}(u_1)\mathbf{A}(u_2)\cdots\mathbf{A}(u_n)\|^2 \\
&\le \sum_{n=0}^{\infty} \int_{z_0}^{z} du_1 \int_{z_0}^{u_1} du_2 \cdots \int_{z_0}^{u_{n-1}} du_n \; K^n \\
&\le \sum_{n=0}^{\infty} \frac{1}{n!}(z-z_0)^n K^n \\
&\le \exp\left[K(z-z_0)\right].
\end{aligned}
\tag{8.2.13}
$$

Since the norm is bounded by a convergent series, it converges and so must the matrix series itself. We conclude therefore that the series in eq.(8.2.9) does indeed converge for any ordinary point z' as was required to be shown. □

As mentioned in the introduction, an important consequence immediately follows:

Corrolary 8.1. *Solutions to the O.D.E. (8.1.1) exist.*

Together with the results of section 3.3 this completes the proof that the space of solutions to eq. (8.1.1) is two-dimensional.

Unfortunately the solution furnished by eqs. (8.2.6) and (8.2.9) is not particularly convenient for practical purposes. A more useful method for generating solutions is to directly substitute a Taylor series solution into eq. (8.1.1), as is illustrated by the following

EXAMPLE:

If z_0 is an ordinary point of the O.D.E. (8.1.1) and $u = z - z_0$ is the displacement of z relative to z_0, then the coefficient functions $p(z)$ and $q(z)$ admit the following Taylor expansion in a neighbourhood of z_0:

$$
p(z) = \sum_{n=0}^{\infty} p_n u^n; \qquad \text{and} \qquad q(z) = \sum_{n=0}^{\infty} q_n u^n. \tag{8.2.14}
$$

Since, by the previous theorem, the solution must also be analytic at z_0 it must similarly admit the expansion:

$$
y(z) = \sum_{n=0}^{\infty} y_n u^n. \tag{8.2.15}
$$

The unknown coefficients of this last expansion can be found in terms of the known quantities p_n and q_n by substituting (8.2.14) and (8.2.15) into the O.D.E. (8.1.1). Using:

$$\begin{aligned} y''(z) &= \sum_{n=0}^{\infty} y_{n+2}(n+1)(n+2)u^n \\ p(z)y'(z) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} p_n y_{m+1}(m+1)u^{n+m} \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} p_k y_{n-k+1}(n-k+1)u^n \\ q(z)y(z) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} q_n y_m u^{n+m} \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} q_k y_{n-k} u^n \end{aligned} \tag{8.2.16}$$

in the O.D.E. allows it to be written as the vanishing of the following series:

$$\sum_{n=0}^{\infty}\left[(n+1)(n+2)y_{n+2} + \sum_{k=0}^{n}\left((n-k+1)p_k y_{n-k+1} + q_k y_{n-k}\right)\right]u^n = 0. \tag{8.2.17}$$

The vanishing of this series is equivalent to the separate vanishing of the coefficient of u^n for each n. This generates the following set of *recursion relations* for the unknown y_n that are equivalent to the original O.D.E.:

$$\begin{aligned} n=0: &\qquad 2y_2 + p_0 y_1 + q_0 y_0 = 0 \\ n=1: &\qquad 6y_3 + 2p_0 y_2 + (q_0+p_1)y_1 + q_1 y_0 = 0 \\ \text{general } n: &\qquad (n+1)(n+2)y_{n+2} + \sum_{k=0}^{n}[(n-k+1)p_k y_{n-k+1} + q_k y_{n-k}] = 0. \end{aligned} \tag{8.2.18}$$

For any $n \geq 0$ the n-th equation clearly determines the coefficient y_{n+2} in terms of the p_k, q_k, and the y_m with $m < n+2$. The two coefficients y_0 and y_1 are *not* determined in this way, however, and are completely undetermined by the O.D.E.. Once y_0 and y_1 are specified all of the remaining coefficients are completely determined by solving the recursion relations (8.2.18). This freedom to choose the first two coefficients comes as no surprise having, as it does, a simple interpretation. To see this notice that the recursion relations are linear in the y_n's and so their solutions have the general form:

$$y_n = A_n[p_k,q_k]y_0 + B_n[p_k,q_k]y_1, \qquad \text{for } n \geq 2. \tag{8.2.19}$$

As indicated in eq. (8.2.19) the quantities A_n and B_n are known functions of the p_k's and q_k's. The solution (8.2.15) to the O.D.E. therefore becomes:

$$y(z) = y_0 \sum_{n=0}^{\infty} A_n[p_k,q_k]u^n + y_1 \sum_{n=0}^{\infty} B_n[p_k,q_k]u^n \tag{8.2.20}$$

which has the form of a linear combination of two basis functions, $w_1(z) = \sum_n A_n(z-z_0)^n$ and $w_2(z) = \sum_n B_n(z-z_0)^n$. ($w_1(z)$ and $w_2(z)$ as defined are clearly linearly independent since w_1 has a Taylor series starting with the term independent of $(z-z_0)$ while w_2 has a Taylor series that leads off with a term linear in $(z-z_0)$. They therefore cannot be proportional to one another as would have to be the case if they were linearly dependent.) The undetermined constants y_0 and y_1 are therefore the two constants of integration that we expect on general grounds to parameterize the two-dimensional space of solutions to the given O.D.E..

This example brings out an important practical consideration in using this technique to generate solutions to second-order O.D.E.'s. The recursion relations (8.2.18) become in general more and more complicated to solve in the form of eq. (8.2.19) because the equation for y_n depends on the values of *all* of the coefficients y_k for $k < n$. In practice the solution for the general coefficient y_n in terms of the p_k's and q_k's is only possible if the equation for each y_n depends on only the previous few y_k's, such as y_{n-1} and y_{n-2}. Inspection of eqs. (8.2.18) shows that this will be so only if the Taylor series for $p(z)$ and $q(z)$ about $z = z_0$ terminates after a small number of terms.

This is unfortunately often not the case. Consider for example Bessel's and Legendre's equations encountered in chapter 5:

$$y''(z) + \frac{1}{z}y'(z) + \left(1 - \frac{n^2}{z^2}\right) y(z) = 0 \qquad \text{(Bessel)} \tag{8.2.21}$$

and:

$$y''(z) - \frac{2z}{1-z^2}y'(z) + \left[\frac{\ell(\ell+1)}{1-z^2} - \frac{m^2}{(1-z^2)^2}\right] y(z) = 0. \qquad \text{(Legendre)} \tag{8.2.22}$$

For Bessel's equation, (5.2.9), $p(z)$ and $q(z)$ have a simple *Laurent* expansion about $z = 0$, however this is not an ordinary point of the O.D.E.. Their *Taylor* expansion about any ordinary point does not terminate. Similarly the associated Legendre equation, (5.3.10), has a simple Laurent expansion about the singular points $z = \pm 1$ but is not particularly simple about any ordinary points.

It is therefore natural to ask whether it is possible to generate a series solution to an O.D.E. by expanding about a singular point.

8.3 Singular Points

According to the definition in the preceding section, a singular point of the O.D.E. (8.1.1) is a point for which either or both of $p(z)$ and $q(z)$ are not analytic.

Definition 8.4. *Isolated Singular Point of an O.D.E.: An isolated singular point of an O.D.E. is a point at which one or both of the functions $p(z)$ and $q(z)$ have an isolated singularity.*

PERVERSITY OF NOTATION:

In all of what follows we assume that the coefficient functions $p(z)$ and $q(z)$ have only isolated singularities. These will be liberally referred to throughout the remainder of the text simply as 'singular points', without the explicit adjective 'isolated'.

We want to be able to generate series solutions to O.D.E.'s by expanding about their singular points. The first question to be addressed therefore is what the analytic behaviour is possible of solutions near a singular point of the O.D.E.. As might be expected, things are more complicated than they are in the neighbourhood of an ordinary point. The remainder of this section is devoted to exploring just how singular solutions can be near a singular point of an O.D.E.. The conclusion that is obtained is that solutions can have one of the following five types of behaviour near a singular point:

1. The solution may be analytic even in the neighbourhood of a singular point. As an example consider the following O.D.E.:
$$y''(z) - \frac{s(s-1)}{z^2} y(z) = 0 \tag{8.3.1}$$
with solution:
$$y(z) = z^s. \tag{8.3.2}$$
If the constant s is a positive integer that is greater than one, then even though the O.D.E. (8.3.1) has $z = 0$ as a singular point, one of its solutions, (8.3.2), is nonetheless analytic there.

2. The solution may have a pole at a singular point. As an example consider again eq. (8.3.1) but this time choose the parameter s to be a negative integer. Then O.D.E. (8.3.1) still has a singular point at $z = 0$, but the solution (8.3.2) now has a simple pole there.

3. The O.D.E. (8.3.1) contains still another example if s is chosen to not be an integer. In this case the solution (8.3.2) exhibits a branch point at $z = 0$ of the form z^s. The discontinuity in the solution across the branch cut is seen to be multiplicative: $y(e^{2\pi i} z) = e^{2\pi i s} y(z)$.

4. The solution could have an isolated singularity that is not a pole: that is it may have an essential singularity at the singular point of the O.D.E.. For an example take the O.D.E.:
$$y''(z) + \frac{2}{z} y'(z) - \frac{1}{z^4} y(z) = 0 \tag{8.3.3}$$
a solution to which is given by:
$$y(z) = \exp\left(\frac{1}{z}\right). \tag{8.3.4}$$
The solution obviously has an essential singularity at $z = 0$ although the coefficient functions of the O.D.E. itself only have poles there.

5. Finally, the solution may have a logarithmic branch point at a singular point. That is, the O.D.E.:
$$y''(z) + \frac{1}{z} y'(z) = 0 \tag{8.3.5}$$
has as solution:
$$y(z) = \log(z). \tag{8.3.6}$$
Along the branch cut this has the additive discontinuity $y(e^{2\pi i} z) = y(z) + 2\pi i$.

It is clearly true that a solution may be (but need not be) singular at a singular point of the O.D.E., and the singularity may be more severe (such as a branch point) than is the singularity in the coefficient functions $p(z)$ and $q(z)$. Even though $p(z)$ and $q(z)$ can be represented by Laurent series around the singular point, it is not in general true that the solutions admit such an expansion.

The next point to be argued is that the examples outlined above completely exhaust the kind of singularities that solutions exhibit at a singular point of an O.D.E..

Theorem 8.2. *If z_0 is an isolated singular point of the O.D.E. (8.1.1) then at least one of the solutions admits the uniformly convergent series representation:*
$$y_1(z) = (z - z_0)^s \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n. \tag{8.3.7}$$
In this equation s represents some complex number. In some cases the second, linearly independent, solution also admits an expansion of the form of eq. (8.3.7). If not the most general form the second solution may take is:
$$y_2(z) = t y_1(z) \log(z - z_0) + (z - z_0)^s \sum_{n=-\infty}^{\infty} b_n (z - z_0)^n. \tag{8.3.8}$$
Again t is a complex number and s is the same complex number that appears in the representation (8.3.7) of the first solution, $y_1(z)$.

The content of this theorem can be summarized as the statement that although the solutions need not be well enough behaved to allow a Laurent series about the singular point, z_0, things are never too much worse.

Proof. (Rather, Sketch of Proof): Consider a disc, D, centred on the singular point, z_0, throughout which the functions $p(z)$ and $q(z)$ are analytic everywhere except for $z = z_0$ itself. Consider also a smaller disc, D', contained within D that does not include the point z_0. We know that the O.D.E. (8.1.1) has two linearly independent solutions, $y_1(z)$ and $y_2(z)$ say, that must be analytic throughout all of D'. Now construct a set of discs, all of the same radius as D', all lying the same distance from z_0 as does D', and all overlapping with both of its nearest neighbours. Using the technique of analytic continuation by power series, for instance, the

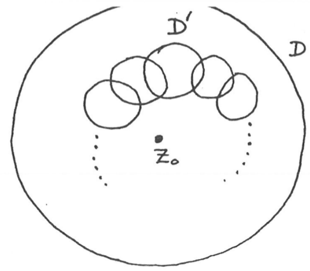

Figure 8.1. The Discs D and D'

two independent solutions can be analytically continued once around z_0 starting from D' and ending up again at D'. In terms of the variable $u = z - z_0$ the analytic continuation of the two solutions in D' once around z_0 can be written as $y_1(e^{2\pi i}u)$ and $y_2(e^{2\pi i}u)$.

The idea then is to compare these continued solutions to the original ones, $y_1(u)$ and $y_2(u)$, inside D'. The point of this comparison is that it reflects the analytic structure of the solutions at the singular point, $z = z_0$. If, for example, we can show that the y_k's are single-valued, *i.e.* $y_k(e^{2\pi i}u) = y_k(u)$, then they must have only an isolated singularity at $z = z_0$ and so must admit a Laurent expansion. If on the other hand they are not continuous when rotating about $z = z_0$, so $y_k(e^{2\pi i}u) \neq y_k(u)$, then the nature of the discontinuity gives information about the type of singularity appearing at $z = z_0$.

We now determine what information the differential equation implies for this analytic continuation. It is clear that, since $p(z)$ and $q(z)$ are analytic everywhere in D except for z_0, that the analytic continuation of a solution to the O.D.E. is also a nonzero solution to the O.D.E.. It follows that the functions $y_k(e^{2\pi i}u) = y_k(u)$ are solutions to the O.D.E. within D' and so can be expanded in terms of the original basis of solutions, $y_k(u)$:

$$\begin{pmatrix} y_1(e^{2\pi i}u) \\ y_2(e^{2\pi i}u) \end{pmatrix} = \mathbf{C} \begin{pmatrix} y_1(u) \\ y_2(u) \end{pmatrix} . \tag{8.3.9}$$

Here $\mathbf{C}$ is a constant, nonvanishing two-by-two matrix called the *analytic continuation matrix.*

Suppose the eigenvalues of this matrix, λ_1 and λ_2, are distinct. (The case where the eigenvalues are degenerate will be considered separately later.) It is a theorem of linear algebra that any such two-by-two matrix may be diagonalized by a similarity transformation. That is, there exists an invertible two-by-two matrix, $\mathbf{S}$, satisfying:

$$\mathbf{SCS}^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{8.3.10}$$

It is convenient, therefore, to consider the following basis of solutions to the O.D.E.:

$$\begin{pmatrix} w_1(u) \\ w_2(u) \end{pmatrix} \equiv \mathbf{S} \begin{pmatrix} y_1(u) \\ y_2(u) \end{pmatrix} \tag{8.3.11}$$

since these have a simpler behaviour once analytically continued around z_0:

$$\begin{aligned}\begin{pmatrix} w_1(e^{2\pi i}u) \\ w_2(e^{2\pi i}u) \end{pmatrix} &= \mathbf{S} \begin{pmatrix} y_1(e^{2\pi i}u) \\ y_2(e^{2\pi i}u) \end{pmatrix} \\ &= \mathbf{SC} \begin{pmatrix} y_1(u) \\ y_2(u) \end{pmatrix} \\ &= \mathbf{SCS}^{-1} \begin{pmatrix} w_1(u) \\ w_2(u) \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} w_1(u) \\ w_2(u) \end{pmatrix} . \end{aligned} \tag{8.3.12}$$

Defining the complex constants s_1 and s_2 by $\lambda_k = e^{2\pi i s_k}$, we see that the solutions $w_k(u)$ have the simple property that:

$$w_k(e^{2\pi i}u) = e^{2\pi i s_k} w_k(u). \tag{8.3.13}$$

An immediate consequence, then, is that the combination $f_k(u) = u^{-s_k} w_k(u)$ is single-valued, $f_k(e^{2\pi i}u) = f_k(u)$, and so has only an isolated singularity at z_0 and therefore admits a Laurent expansion centred at z_0 throughout D.

We have proven that if $\lambda_1 \neq \lambda_2$ then it is always possible to construct a basis of solutions to the O.D.E. of the following form:

$$w(z) = (z - z_0)^s \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \tag{8.3.14}$$

about an isolated singular point. This is precisely the form given in eq. (8.3.7). The eigenvalues of the analytic continuation matrix $\mathbf{C}$ are related to the parameter s by $\lambda = e^{2\pi i s}$.

The remaining case to consider is when both eigenvalues of the analytic continuation matrix, $\mathbf{C}$, are equal: $\lambda_1 = \lambda_2 \equiv \lambda$. This is equivalent to the requirement that $s_2 - s_1 = n$ where n is an integer. Then $\mathbf{C}$ may or may not be diagonalizable by a similarity transformation of the form (8.3.10). If so, then $\mathbf{C}$ must be proportional to the unit matrix and so is already diagonal. All solutions to the differential equation in this case must must satisfy eq. (8.3.7) under analytic continuation around z_0 with s given by $\lambda = e^{2\pi i s}$.

The more interesting examples are when the eigenvalues are equal but the matrix is not diagonalizable. It is always possible to put any such two-by-two matrix into *Jordan canonical form* by a similarity transformation. *i.e.*:

$$\mathbf{SCS}^{-1} = \begin{pmatrix} \lambda & 0 \\ \mu & \lambda \end{pmatrix} . \tag{8.3.15}$$

μ here is some complex number. If $\mu = 0$ then the matrix $\mathbf{C}$ is obviously diagonalizable and the discussion reduces to that presented in the previous paragraph.

As before define a new basis of solutions, $w_k(u)$, to the O.D.E. in terms of the original ones, $y_k(u)$, by eq. (8.3.11). These satisfy:

$$\begin{pmatrix} w_1(e^{2\pi i}u) \\ w_2(e^{2\pi i}u) \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ \mu & \lambda \end{pmatrix} \begin{pmatrix} w_1(u) \\ w_2(u) \end{pmatrix} . \tag{8.3.16}$$

This defines the transformation properties of the solutions $w_k(u)$ under analytic continuation about $z = z_0$. Eq. (8.3.16) can be re-expressed by constructing two functions,

$$f_1(u) = u^{-s} w_1(u) \tag{8.3.17}$$

and

$$f_2(u) = u^{-s}[w_2(u) - t w_1(u) \log u]. \tag{8.3.18}$$

The complex parameters s and t arising here are defined by: $\lambda = e^{2\pi i s}$ and $\mu = 2\pi i \lambda t$. The functions $f_k(u)$ so constructed are single-valued in D', $f_k(e^{2\pi i}u) = f_k(u)$, and so can be represented by a Laurent series throughout D. By inverting the relations (8.3.17) and (8.3.18) for the $w_k(u)$'s in terms of the $f_k(u)$'s we find that $w_1(u)$ admits the expansion (8.3.7) and $w_2(u)$ can be written in the form (8.3.8).

This is what was required to be shown. □

Notice that it is always possible to construct at least one solution of the O.D.E. with the series form (8.3.7) about an isolated singular point. The coefficients in this series together with the power s appearing in the prefactor $(z - z_0)^s$ may be in principle determined in the same manner as were the coefficients in an expansion about an ordinary point: by directly inserting the series into the O.D.E.. As we shall see the result is not of much practical use in explicitly generating solutions.

EXAMPLE:

Suppose z_0 is an isolated singular point of the O.D.E. (8.1.1). Then the coefficient functions have Laurent expansions:

$$p(z) = \sum_{-\infty}^{\infty} p_n u^n \qquad \text{and} \qquad q(z) = \sum_{-\infty}^{\infty} q_n u^n. \tag{8.3.19}$$

u is as usual defined as the difference $u = z - z_0$. By the previous theorem we are guaranteed the existence of a solution with a uniformly convergent power-series expansion:

$$y(z) = u^s \sum_{-\infty}^{\infty} y_n u^n \tag{8.3.20}$$

Plugging all of this into the differential equation gives:

$$\sum_{n=-\infty}^{\infty} \Bigg[(s+n+1)(s+n+2) y_{n+2} + \sum_{k=-\infty}^{\infty} \big((s+n-k+1) p_k y_{n-k+1} + q_k y_{n-k} \big) \Bigg] u^{s+n} = 0. \tag{8.3.21}$$

The recursion relation for the y_n's is found by setting the coefficient of each power of u separately to zero. The result is:

$$(s+n+1)(s+n+2)y_{n+2} + \sum_{k=-\infty}^{\infty} ((s+n-k+1)p_k y_{n-k+1} + q_k y_{n-k}) = 0. \qquad (8.3.22)$$

This set of infinitely many coupled equations determines, in principle, the unknown y_n's as well as the parameter s. This is obviously not of much practical use since in general all of the equations simulataneously involve all of the unknowns.

This pessimistic conclusion could be avoided if eqs. (8.3.22) were to have solutions for which the Laurent series should terminate in either the direction of increasing or decreasing powers of $u = (z - z_0)$. It suffices to consider the case where the series terminates for negative powers of u because the other case reduces to this one after the change of variables $u \to u^{-1}$. It would then be possible to proceed as for an ordinary point and generate all of the y_n's in terms of the first few. We are led, then, to ask how singular the coefficient functions $p(z)$ and $q(z)$ may be at $z = z_0$ without introducing an essential singularity into the solution.

8.4 Regular Singular Points

With the lessons of the previous example in mind, suppose that $p(z)$ has a pole of order m at $z = z_0$ and that $q(z)$ has a pole there of order n. Their Laurent expansions about $z = z_0$ become:

$$\begin{aligned} p(u) &= u^{-m}[p_0 + p_1 u + ...]; \qquad p_0 \neq 0 \\ q(u) &= u^{-n}[q_0 + q_1 u + ...]; \qquad q_0 \neq 0. \end{aligned} \qquad (8.4.1)$$

We are looking for the existence of solutions of the form of (8.3.7) with the restriction that the series terminates for sufficiently small powers of u:

$$y(u) = u^s[y_0 + y_1 u + ...]; \qquad y_0 \neq 0. \qquad (8.4.2)$$

To answer this question substitute eqs. (8.4.1) and (8.4.2) into the O.D.E. (8.1.1). The result is:

$$\begin{aligned} &\{s(s-1)y_0 u^{s-2} + s(s+1)y_1 u^{s-1} + ...\}+ \\ +&\{sp_0 y_0 u^{s-m-1} + [(s+1)p_0 y_1 + sp_1 y_0]u^{s-m} + ...\}+ \\ +&\{q_0 y_0 u^{s-n} + [q_0 y_1 + q_1 y_0]u^{s-n+1} + ...\} = 0. \end{aligned} \qquad (8.4.3)$$

The first curly bracket is the contribution of $y''(z)$, the second is the $p(z)y'(z)$ part and the last is the $q(z)y(z)$ term. To satisfy the O.D.E. we must set the coefficient of each power of u to zero. There are three cases to consider:

1. $n > \max(2, m+1)$: If n is greater than the larger of 2 and $m+1$, then there is no solution to the O.D.E. that satisfies the ansatz of eq. (8.4.2). This may be seen by

realizing that in this case the lowest power of u that appears in eq. (8.4.3) is u^{s-n}. Since the coefficient of this term is $q_0 y_0$ which does not vanish by assumption, there is no way to satisfy eq. (8.4.3).

2. $m > \max(n-2, 1)$: In the event that $m+1$ is greater than both $n-1$ and 2 the lowest power of u is u^{s-m-1}. The coefficient of this term is either $sp_0 y_0$ or $(sp_0 + q_0)y_0$ depending on whether or not $m+1 = n$. The vanishing of this coefficient determines s to be zero or $-q_0/p_0$, respectively. This equation, determining as it does the value of s, is called the *indicial equation*. The vanishing of the coefficient of the next-lowest power, u^{s-m}, then determines y_1 and so on.

   It would appear that in this case a solution of the form (8.4.2) would exist but this is deceptive. The problem is that although the recursion relation (8.4.3) generates values for s and the coefficients y_k the resulting series need not converge for any nonzero value of u. As an example consider the O.D.E. for which $m = n = 2$:

$$y''(z) - \frac{1}{z^2}y'(z) + \frac{1}{z^2}y(z) = 0. \tag{8.4.4}$$

   The indicial and recursion equations become for this example:

$$\begin{aligned} s &= 0 \\ y_{k+1} &= \left[\frac{k(k-1)+1}{k+1}\right] y_k \end{aligned} \tag{8.4.5}$$

   In order to test the resulting series for convergence consider the ratio of successive terms:

$$\frac{y_{k+1}z^{k+1}}{y_k z^k} = \left[\frac{k(k-1)+1}{k+1}\right] z. \tag{8.4.6}$$

   The limit of this ratio as $k \to \infty$ diverges for any nonzero z. By the ratio test the corresponding series must have zero radius of convergence. The series is an asymptotic series rather than a convergent one. It follows that we are not guaranteed a solution of the form (8.4.2) in this case, even though it is possible to generate a solution to the recursion relation and indicial equation.

3. $n \leq 2$ and $m \leq 1$: The only remaining case is that for which $p(z)$ has at most a simple pole and $q(z)$ has at most a double pole at $z = z_0$. It is convenient to take, in the general formula (8.4.3), $n = 2$ and $m = 1$. This really includes the more general cases if the possibility is allowed that q_0 and p_0 may be zero. In fact, if all of p_0, q_0 and q_1 vanish we retrieve the case where z_0 is an ordinary point. The lowest power of u in eq. (8.4.3) is in this case u^{s-2} and the vanishing of the coefficient of this term imposes a quadratic condition apon s. As in the previous case this indicial equation and the recursion relations may be solved in principle to generate a series of the form (8.4.2). The question again arises whether the resulting series converges or is merely

asymptotic to a solution to the O.D.E.. The fundamental claim to be made here now is that for at least one solution to the O.D.E. this series converges. We return to the explicit construction of such solutions in more detail below once this result has been more formally stated.

Case (3) above motivates the following

Definition 8.5. *Regular Singular Point: If z_0 is an isolated singular point of the O.D.E. (8.1.1) for which the coefficient function $p(z)$ has at most a simple pole and $q(z)$ has at most a double pole, then z_0 is called a* regular singular point *of the O.D.E..*

Definition 8.6. *Irregular Singular Point: An* irregular singular point *of an O.D.E. is an isolated singular point that is not regular.*

The basic statement about the convergence of the series solution (8.4.2) about a regular singular point is contained in the following theorem that is stated here without proof. It may be proven by a slight generalization of the argument, given in section 8.2, that establishes that solutions are analytic at an ordinary point.

Theorem 8.3. *Fuch's Theorem: The series form (8.3.7) for one of the solutions to the O.D.E. $y'' + py' + qy = 0$ always terminates for sufficiently large negative n. That is, the solution can be written as a convergent expansion of the form:*

$$y(z) = (z - z_0)^s \sum_{n=0}^{\infty} y_n (z - z_0)^n \tag{8.4.7}$$

about a regular singular point, z_0. The radius of convergence of this expansion is at least as large as the smaller of the two radii of convergence of the Laurent series for the functions $p(z)$ and $q(z)$ about z_0.

We turn now to the construction of this solution.

EXAMPLE:

Suppose z_0 is a regular singular point of the differential equation (8.1.1). Then the functions $up(u)$ and $u^2 q(u)$ are analytic at $u = z - z_0 = 0$ and so admit the Taylor expansion:

$$\begin{aligned} up(u) &= \sum_{n=0}^{\infty} p_n u^n \\ u^2 q(u) &= \sum_{n=0}^{\infty} q_n u^n. \end{aligned} \tag{8.4.8}$$

Fuchs's theorem ensures the existence of a solution of the form:

$$y(u) = \sum_{n=0}^{\infty} y_n u^{n+s}. \tag{8.4.9}$$

The differential equation becomes:

$$\sum_{n=0}^{\infty}\left[(s+n)(s+n-1)y_n+\sum_{k=0}^{n}((s+n-k)p_k y_{n-k}+q_k y_{n-k})\right]u^{s+n}=0. \tag{8.4.10}$$

The $n=0$ term gives rise to the indicial equation:

$$I(s)\equiv s(s-1)+sp_0+q_0=0 \tag{8.4.11}$$

and the remaining terms provide the recursion relation:

$$I(s+n)y_n=F_n[s;y_0,...,y_{n-1}] \qquad n\geq 1. \tag{8.4.12}$$

In this last equation $I(s)$ is the function defined by eq. (8.4.11) and $F_n[s;y_0,...,y_{n-1}]$ is given explicitly by:

$$F_n[s;y_0,...,y_{n-1}]=\sum_{k=1}^{n}((s+n-k)p_k+q_k)\,y_{n-k}. \tag{8.4.13}$$

Consider now the solutions to these conditions. The indicial equation, being quadratic in s, has two roots s_1 and s_2. Each of these roots may be substituted into the recursion relations (8.4.12) to generate the coefficients:

$$y_n(s_1)=\frac{F_n[s_1;y_0,...,y_{n-1}]}{I(s_1+n)}; \qquad n\geq 1 \tag{8.4.14}$$

and:

$$y_n(s_2)=\frac{F_n[s_2;y_0,...,y_{n-1}]}{I(s_2+n)}; \qquad n\geq 1 \tag{8.4.15}$$

Notice that, unlike the situation encountered when expanding about an ordinary point, in general only y_0 is left undetermined by these equations. This implies that each root of the indicial equation generates only a single solution to the O.D.E.. If s_2-s_1 is not an integer, then the two solutions so generated behave differently as $z\to z_0$ and so cannot be proportional to one another. This implies that they must be linearly independent and so form a basis for the space of solutions.

8.5 When the Method Fails: $s_2-s_1=N$

In the case that the difference between the roots s_1 and s_2 is an integer it need no longer be true that both solutions to the differential equation have the form (8.4.9). This is a reflection of a property that has already been encountered. To see this notice that comparison of the ansatz (8.4.9) to the general form (8.3.12) shows that the roots s_k are related to the eigenvalues λ_k of the analytic continuation matrix (introduced in section 8.3) by $\lambda_k=e^{2\pi i s_k}$. Consequently the condition that s_2 and s_1 differ by an integer is equivalent to the condition that the two eigenvalues be equal: $\lambda_1=\lambda_2$. This is precisely the condition determined in section 8.3 under which the second solution may contain a logarithmic singularity, for which the expression (8.3.12) fails.

The remainder of this section is devoted to a detailed exploration of what goes wrong with the solutions (8.4.14) and (8.4.15) when $s_2 = s_1 + N$, for integer N. This is followed with an alternate method of explicitly constructing the second solutions in this case. The discussion proceeds slightly differently depending on whether or not N vanishes, so we consider these two possibilities in turn.

1. $s_2 = s_1 \equiv \alpha$: Suppose the indicial equation has the form:

$$I(s) = (s - \alpha)^2 \tag{8.5.1}$$

with degenerate root $s = \alpha$. In this case the recursion relations (8.4.14) and (8.4.15) both have solutions but these solutions are not distinct since they both reduce to:

$$y_n(\alpha) = \frac{F_n[\alpha; y_0, ..., y_{n-1}]}{n^2}; \qquad n \geq 1. \tag{8.5.2}$$

2. $s_2 - s_1 = N$: In the event that the roots of the indicial equation differ by an integer it is always possible to label the roots $s_2 = \alpha$ and $s_1 = \alpha - N$, with N positive. The indicial equation in this notation is

$$I(s) = (s - \alpha)(s - \alpha + N). \tag{8.5.3}$$

In this case the recursion relation (8.4.15) using the root α is perfectly well-behaved:

$$y_n(\alpha) = \frac{F_n[\alpha; y_0, ..., y_{n-1}]}{n(n + N)}; \qquad n \geq 1. \tag{8.5.4}$$

A potential difficulty arises however when the other root is used in eq. (8.4.14):

$$y_n(\alpha - N) = \frac{F_n[\alpha - N; y_0, ..., y_{n-1}]}{n(n - N)}; \qquad n \geq 1. \tag{8.5.5}$$

The difficulty lies in the term with $n = N$ for which the denominator of eq. (8.5.5) vanishes. There are two possibilities. If $F_N[\alpha - N; y_0, ..., y_{N-1}]$ should also vanish, then the differential equation (8.4.12) is satisfied for any y_N. The polynomial in u formed by taking $y_N = 0$ is therefore a solution of the O.D.E.. The remainder of the coefficients $y_{N+k}(\alpha - N)$ defined by the recursion relation eq. (8.5.5) then generate the same series solution as does $y_k(\alpha)$ and so need not be separately considered. This implies that both solutions have the form (8.4.9) and y_0 and y_N are the free constants of integration.

If, on the other hand, $F_N[\alpha - N; y_0, ..., y_{N-1}]$ is not zero, then it is impossible to satisfy the recursion relation for any choice of y_N and so the second solution to the O.D.E. cannot be written in the form (8.4.9).

All is not lost, however, since an alternative method is available with which to generate the second solution in these degenerate cases. This method supplements the general technique presented in section 3.3 for generating second solutions and is more useful when the first solution is only known in series form. Since the steps differ for the two cases where $s_1 = s_2$ or where s_1 and s_2 differ by a nonzero integer, these cases are again handled separately.

1. $s_1 = s_2 = \alpha$: It is convenient for the purposes of this argument to denote the O.D.E. (8.1.1) as:

$$\mathcal{L}y(z) = 0 \tag{8.5.6}$$

with

$$\mathcal{L} = \frac{d^2}{dz^2} + p(z)\frac{d}{dz} + q(z). \tag{8.5.7}$$

Consider, then, the function $Y(s;z)$ defined by the series:

$$Y(s;z) = (z - z_0)^s \sum_{n=0}^{\infty} y_n(s)(z - z_0)^n \tag{8.5.8}$$

in which the s-dependent coefficients, $y_n(s)$, satisfy the recursion relation (8.4.12) but for which s is arbitrary and does *not* satisfy the indicial equation (8.4.11). Clearly evaluation of this function at $s = \alpha$ produces the known series solution to the O.D.E.: $y_1(z) = Y(\alpha; z)$.

The plan is to present a construction of the second solution using this function $Y(s;z)$. It is important to realize that $Y(s;z)$ itself is not a solution to the O.D.E. for any value of s except α. To see this explicitly apply the operator $\mathcal{L}$ on $Y(s;z)$. Inserting the series (8.5.8) into the differential equation gives eq. (8.4.10) all the terms of which vanish by virtue of the recursion relation, *with the exception of the first*:

$$\mathcal{L}Y(s;z) = I(s)(z - z_0)^s = (s - \alpha)^2(z - z_0)^s. \tag{8.5.9}$$

From eq. (8.5.9) it is clear how to construct a second solution to the O.D.E. (8.5.6): differentiate $Y(s;z)$ once with respect to s and evaluate the result at $s = \alpha$:

$$y_2(z) = \left.\frac{\partial Y(s;z)}{\partial s}\right|_{s=\alpha}. \tag{8.5.10}$$

To see that this is indeed a solution, differentiate eq. (8.5.9) with respect to s and evaluate at $s = \alpha$, using the fact that the operator $\mathcal{L}$ does not depend on s. This gives:

$$\left[\mathcal{L}\left(\frac{\partial Y}{\partial s}\right)\right]_{s=\alpha} = (s - \alpha)(z - z_0)^s[2 + (s - \alpha)\log(z - z_0)]|_{s=\alpha} = 0. \tag{8.5.11}$$

It is furthermore an independent solution since explicit differentiation of $Y(s;z)$ gives:

$$y_2(z) = y_1(z)\log(z - z_0) + \sum_{n=0}^{\infty}\left(\frac{\partial y_n(s)}{\partial s}\right)_{s=\alpha}(z - z_0)^{n+\alpha}. \tag{8.5.12}$$

$y_2(z)$ must be linearly independent of $y_1(z)$ because otherwise they would be proportional to one another. However they cannot be proportional to one another since $y_2(z)$ has a logarithmic singularity as $z \to z_0$ which $y_1(z)$ does not share.

2. $s_2 = \alpha$ and $s_1 = \alpha - N$ for N a positive integer: In this case we must use a slight modification of the same trick. Consider again the function $Y(s;z)$, defined just as before by (8.5.8) in which the recursion relations are imposed but the parameter s is kept arbitrary. The known solution is in this case obtained by evaluating $Y(s;z)$ at $s = \alpha$: $y_1(z) = Y(\alpha; z)$. We are interested in the case where evaluation at $s = \alpha - N$ does not generate the second solution because the recursion relation implies $y_N(\alpha - N)$ is infinite. This shows up in the function $Y(s;z)$ as a simple pole for $y_N(s)$ at $s = \alpha - N$ as is seen from eq. (8.4.12):

$$y_N(s) = \frac{F_N[s; y_0, ..., y_{N-1}]}{I(s+N)} = \frac{F_N[s; y_0, ..., y_{N-1}]}{(s-\alpha+N)(s-\alpha+2N)}. \tag{8.5.13}$$

In this case the action of the operator $\mathcal{L}$ on $Y(s;z)$ is:

$$\mathcal{L}Y(s;z) = I(s)(z-z_0)^s = (s-\alpha)(s-\alpha+N)(z-z_0)^s. \tag{8.5.14}$$

Clearly the second solution is not given by simply differentiating $Y(s;z)$ with respect to s and evaluating at one of the roots because eq. (8.5.14) implies that this does not satisfy the O.D.E.. Instead take:

$$y_2(z) = \frac{\partial}{\partial s}[(s-\alpha+N)Y(s;z)]\bigg|_{s=\alpha-N}. \tag{8.5.15}$$

This satisfies the O.D.E. since multiplication of eq. (8.5.14) by $(s-\alpha+N)$ followed by differentiation with respect to s and evaluation at $\alpha - N$ gives zero. It is also finite as $s \to \alpha - N$ since the pole has been cancelled. Finally, it is linearly independent since unlike $y_1(z)$ it exhibits a logarithmic singularity as $z \to z_0$.

Notice that the alternative procedure of multiplying by $(s-\alpha)$, differentiating and evaluating at $s = \alpha$ does not give a new solution:

$$\frac{\partial}{\partial s}[(s-\alpha)Y(s;z)]\bigg|_{s=\alpha} = Y(\alpha;z) = y_1(z). \tag{8.5.16}$$

The first equality follows from the observation that none of the terms in $Y(s;z)$ are singular as $s \to \alpha$ so the only contribution comes when the factor $(s-\alpha)$ itself is differentiated.

This concludes our discussion about the construction of explicit solutions to second-order linear O.D.E.'s.

9 Classification of Ordinary Differential Equations

9.1 Introduction

From the previous chapter we have seen that the solutions to essentially any linear second-order O.D.E. can be constructed by searching for power-series solutions about any ordinary or a regular singular point. The key observation to be made may be stated as the following 'folk theorem':

FOLK THEOREM:

All of the O.D.E.'s that generally arise in mathematical physics fall into two categories: 1) those with no irregular singular points and three or fewer regular singular points, and 2) those with at most one regular and one irregular singular point. Even better: the equations having one regular and one irregular point don't have arbitrarily singular behaviour at the irregular point. In practice all of the irregular singular points found in the O.D.E.'s of interest to traditional mathematical physics can be obtained by the *confluence* of two regular singular points. That is to say they may be obtained from an equation with three regular singular points by allowing two of the regular points to coincide.

The upshot is that only those equations involving three or fewer regular singular points occur in practical problems. It is therefore of some interest to know that all such second-order linear differential equations have been classified once and for all. This chapter is devoted to presenting this classification. The resulting O.D.E.'s are then solved in chapter 10.

Before proceeding with the classification however it is necessary to establish what to do with the point 'at infinity'. This must be settled because the classification theorem gives a complete list of O.D.E.'s with a given number of singular points provided that 'infinity' is included among the ordinary or regular points.

Since the regularity of a point is tied up with the analytic properties of the functions $p(z)$ and $q(z)$ at that point, some definitions concerning the treatment of analytic functions at 'infinity' must be recalled.

Definition 9.1. *A function $f(z)$ is said to be analytic as $z \to \infty$, or analytic at infinity, if after the change of variable $w = 1/z$ it is analytic at $w = 0$.*

This implies that, at the very least, $f(z)$ is bounded as $z \to \infty$ in any direction.

With this definition in mind, the classification of the point at infinity as an ordinary, regular-singular or irregular-singular point of an O.D.E. is done by changing variables $w = 1/z$ and asking the same question at $w = 0$. After this change of variables the O.D.E.:

$$\frac{d^2y}{dz^2} + p(z)\frac{dy}{dz} + q(z)y(z) = 0 \tag{9.1.1}$$

becomes

$$\frac{d^2y}{dw^2} + P(w)\frac{dy}{dw} + Q(w)y(w) = 0 \tag{9.1.2}$$

with the functions $P(w)$ and $Q(w)$ related to the original functions $p(z)$ and $q(z)$ by:

$$P(w) = \frac{2}{w} - \frac{1}{w^2} p\left(\frac{1}{w}\right) \tag{9.1.3}$$

and

$$Q(w) = \frac{1}{w^4} q\left(\frac{1}{w}\right). \tag{9.1.4}$$

Definition 9.2. *$z = \infty$ is an ordinary point of the O.D.E. (9.1.1) if both of the functions $P(w)$ and $Q(w)$ are analytic at $w = 0$.*

Notice that this is *not* the same as the requirement that $p(z)$ and $q(z)$ be analytic at infinity. The conditions that $p(z)$ and $q(z)$ must satisfy are easily found by comparison with eqs. (9.1.3) and (9.1.4). We see that for large z the functions $p(z)$ and $q(z)$ fall to zero (in all directions!) in a very specific way:

$$p(z) = \frac{2}{z} + O(z^{-2}) \qquad \text{as } z \to \infty \tag{9.1.5}$$

and

$$q(z) = O(z^{-4}) \qquad \text{as } z \to \infty. \tag{9.1.6}$$

Notice, in eq. (9.1.5), that the coefficient of z^{-1} in the asymptotic form for $p(z)$ must be exactly 2. It is *not* sufficient that $p(z)$ be merely $O(z^{-1})$.

Definition 9.3. *Infinity is a regular singular point of the O.D.E. (9.1.1) if the function $P(w)$ has at most a simple pole at $w = 0$ and the function $Q(w)$ has at most a double pole at $w = 0$.*

Again, for functions $p(z)$ and $q(z)$ these conditions can be translated, using eqs. (9.1.3) and (9.1.4), into:

$$p(z) = O(z^{-1}) \qquad \text{as } z \to \infty \tag{9.1.7}$$

and

$$q(z) = O(z^{-2}) \qquad \text{as } z \to \infty. \tag{9.1.8}$$

The important point about these definitions is that they greatly constrain how quickly $p(z)$ and $q(z)$ must vanish for large z. We may now turn to the classification theorem for equations with three or fewer regular singular points.

9.2 The Hypergeometric Equation

The principal tools in this analysis are theorems 6.1 and 6.2 of section 6.2. These state that:

1. The only bounded entire functions are the constant functions, and:

2. The only entire functions that grow like z^n for large z are polynomials of degree n.

We now explicitly list all O.D.E.'s with three or fewer regular singular points.

1. Category O: There are no equations with no singular points.

   Proof. The argument that establishes this claim is typical of those used in the cases that follow. We consider all the properties the coefficient functions $p(z)$ and $q(z)$ must have and show that they are incompatible. Their first property is that they must both be entire functions since they cannot have any singular points for any finite z. Next, infinity must also be an ordinary point so eqs. (9.1.5) and (9.1.6) hold. These imply, among other things, that $p(z)$ and $q(z)$ are bounded as $z \to \infty$. Together with the theorem (1) above, it follows that $p(z)$ and $q(z)$ are both constants, independent of z. However this is inconsistent with condition (9.1.5) since this requires that the large-z behaviour of $p(z)$ is $2/z+O(z^{-2})$. There are therefore no such differential equations. □

2. Category I: The only equation with exactly one singular point can be written as:

$$y''(z) + \frac{2}{z}y'(z) = 0. \tag{9.2.1}$$

   Proof. It must be recognized that changes of independent and dependent variables can always be used to put an O.D.E. into a preferred form. This freedom is used extensively in what follows, the more so for the differential equations having more numerous singular points. In the present case suppose that the sole regular singular point is located at z_1. We are free to put this point at the origin, $z = 0$, by the change of variables $u = z - z_1$. (If the singular point should be at infinity it may be moved to the origin *via* the usual choice: $w = 1/z$.) Having done so, it must be that $p(z)$ and $q(z)$ are analytic throughout the complex plane except for the origin where $p(z)$ potentially has a simple pole and $q(z)$ may have a double pole. In other words $zp(z)$ and $z^2q(z)$ must be entire functions. Finally, since infinity is an ordinary point $p(z)$ and $q(z)$ must satisfy eqs. (9.1.5) and (9.1.6). These imply that $zp(z) = 2 + O(z^{-1})$ and $z^2q(z) = O(z^{-2})$ for large z. $zp(z)$ and $z^2q(z)$ are therefore both entire bounded functions and so by theorem (1) must be constants. Comparison with their known limits for large z shows that the only allowed constants are 2 and 0 respectively. It follows that $p(z) = 2/z$ and $q(z) = 0$, which reproduces eq. (9.2.1).

   The general solution to this equation may be written down by inspection:

$$y(z) = A + \frac{B}{z} \tag{9.2.2}$$

   in which A and B are the constants of integration. □

3. Category II: The only equation with exactly two regular singular points is the Euler equation:

$$y''(z) + \frac{a}{z}y'(z) + \frac{b}{z^2}y(z) = 0. \tag{9.2.3}$$

   where a and b are arbitrary complex constants.

Proof. The argument proceeds just as for the preceding examples. Suppose the singular points are located at z_1 and z_2. The change of variables $w = (z - z_1)/(z - z_2)$ then puts the singular points at zero and infinity. With this choice $zp(z)$ and $z^2q(z)$ must be entire functions. At infinity eqs.(9.1.7) and (9.1.8) imply $zp(z) = O(z^0)$ and $z^2q(z) = O(z^0)$, which says that they are bounded by constants as $z \to \infty$. Theorem (1) then ensures that they are constants, a and b say, but the values of these constants are not constrained any further. The conclusion is: $p(z) = a/z$ and $q(z) = b/z^2$ as claimed.

The general solution to this equation is well known:

$$y(z) = Az^{s_1} + Bz^{s_2} \tag{9.2.4}$$

where s_1 and s_2 are the distinct roots of $s(s-1)+as+b = 0$. If the roots of this equation are not distinct, then the two solutions in eq. (9.2.4) become linearly dependent and the general solution is instead:

$$y(z) = Az^s + Bz^s \log z. \tag{9.2.5}$$

$\square$

4. Category III: The most general equation with precisely three regular singular points is the *Hypergeometric Equation*:

$$z(1 - z)y''(z) + [c - (1 + a + b)z]y'(z) - aby(z) = 0. \tag{9.2.6}$$

   a, b and c are complex constants.

Proof. The argument is familiar but more lengthy in this case. If the three singular points are originally at z_1, z_2 and z_3, then the change of variables:

$$w = \left(\frac{z - z_1}{z - z_2}\right)\left(\frac{z_3 - z_2}{z_3 - z_1}\right) \tag{9.2.7}$$

may be used to put them to zero, one and infinity. Once this has been done, the functions $z(z - 1)p(z)$ and $z^2(z - 1)^2q(z)$ must be entire functions. At infinity eqs. (9.1.7) and (9.1.8) imply that $z(z - 1)p(z)$ grows like z and $z^2(z - 1)^2q(z)$ grows like z^2 as $z \to \infty$. Consulting theorem (2) quoted at the beginning of this section then gives:

$$p(z) = \frac{\alpha' + \beta' z}{z(z - 1)} = \frac{\alpha}{z} + \frac{\beta}{z - 1} \tag{9.2.8}$$

and

$$q(z) = \frac{\gamma' + \delta' z + \epsilon' z^2}{z^2(z - 1)^2} = \frac{\gamma}{z(z - 1)} + \frac{\delta}{z^2} + \frac{\epsilon}{(z - 1)^2}. \tag{9.2.9}$$

The Greek letters denote arbitrary constants. The second equality in these two equations follows by re-expressing the given quantity by partial fractions. The relations between the primed and unprimed constants is then seen to be: $\alpha = -\alpha'$, $\beta = \alpha' + \beta'$, $\gamma = -2\gamma' - \delta'$, $\delta = \gamma'$ and $\epsilon = \gamma' + \delta' + \epsilon'$.

This is not the end of the story however, since eqs. (9.2.7) and (9.2.8) are not yet equivalent to eq. (9.2.6). In order to establish the connection we wish to use up the remaining freedom to change variables to further restrict the functions $p(z)$ and $q(z)$. In particular, this freedom will be chosen to set the constants δ and ϵ to zero.

In order to do so, it is necessary to enumerate exactly how much freedom there is to change variables. In general there are two types of variable changes that may be performed. These consist of changes in the independent variable: $u = u(z)$ and changes of the dependent variable: $y = y(v)$. This is slightly deceptive, however, since an arbitrary such change is not really of interest. What is really required is the most general variable change that preserves the linearity of the O.D.E., does not introduce any new singular points, and keeps the singular points at zero, one and infinity. We consider changes of independent and dependent variable in turn.

For changes in the independent variable, the condition that no new singularities be introduced requires that $u = u(z)$ be a one-to-one transformation of the complex plane (including the point at infinity) to itself. A transformation such as $u = z^2$ is, for example, ruled out since its inverse is $z = \sqrt{u}$ which introduces a new branch point into functions like $p[z(u)]$. These conditions are very restrictive. They imply that $u(z)$ vanishes at precisely one point, z_0, and likewise goes to infinity only at a single point, z_∞. Furthermore, the Inverse Function Theorem of calculus implies that both du/dz and dz/du must be nonzero everywhere (including infinity) except possibly for the points z_0 and z_∞. These conditions together with theorem (1) above imply that du/dz is determined to be: $du/dz = C/(z - z_\infty)^2$. Integration then gives as the general solution the set of *fractional-linear* or *homographic* transformations:

$$u(z) = \frac{Az + B}{Cz + D}. \tag{9.2.10}$$

The four constants A, B, C and D are not independent since they may all be multiplied by a factor without changing $u(z)$. We may choose, then, without loss of generality, that they satisfy the condition $AD - BC = 1$. It is the freedom to perform these transformations that has been used this far to place the singular points at zero, one and infinity. If the transformation is required to keep the singular points at these three places, then the remaining freedom to change independent variable reduces to the group of six transformations that permute these three points: $u = z$, $u = 1/z$, $u = 1 - z$, $u = 1/(1 - z)$, $u = 1 - 1/z$ and $u = -z/(1 - z)$. The set of these transformations is often called the *hypergeometric group*. These transformations have some intrinsic interest but are ignored here since they cannot be used to set the two parameters δ and ϵ to zero. This requires instead at least a two-parameter set of transformations. For this we must therefore turn to changes of the dependent variable.

The principal requirement that a change of variable $y = y(v)$ must satisfy is that it preserve the linearity of the O.D.E.. This implies that the change of variables itself must be linear: $y(z) = f(z)v(z)$. The function $f(z)$ is further constrained by the requirement that the change of variables not introduce new singular points into the O.D.E.. The next step is to determine what class of functions $f(z)$ satisfy this condition.

To determine this notice that in the new variables the O.D.E. becomes:

$$v''(z) + P(z)v'(z) + Q(z)v(z) = 0 \tag{9.2.11}$$

with the new coefficient functions related to the old ones by:

$$P(z) = p(z) + 2\frac{f'}{f}(z) \tag{9.2.12}$$

and

$$Q(z) = q(z) + p(z)\frac{f'}{f}(z) + \frac{f''}{f}(z). \tag{9.2.13}$$

The condition implied for $f(z)$ is the requirement that both the functions $p(z)$ and $P(z)$ must satisfy eq. (9.2.8) and that both $q(z)$ and $Q(z)$ must satisfy eq. (9.2.9). Comparison of eq. (9.2.12) with eq. (9.2.8) shows that $f(z)$ must satisfy:

$$\frac{f'}{f}(z) = \frac{A}{z} + \frac{B}{z-1}. \tag{9.2.14}$$

A and B are arbitrary complex parameters. Eq. (9.2.14) has as its solution

$$f(z) = z^A(z-1)^B. \tag{9.2.15}$$

An irrelevant multiplicative constant has been dropped in this equation. It is easily verified that eq. (9.2.9) imposes no new conditions on $f(z)$, so (9.2.15) is the most general change of variables that preserves the number and type of singular points.

The free parameters A and B may now be chosen to put the differential equation into the standard Hypergeometric form. The conventional choice for this standard form is to use A and B to set the parameters δ and ϵ to zero. To see what this implies substitute eqs. (9.2.12), (9.2.13) and (9.2.15) into eqs. (9.2.8) and (9.2.9):

$$P(z) = \frac{\alpha + 2A}{z} + \frac{\beta + 2B}{z-1} \tag{9.2.16}$$

and

$$Q(z) = \frac{\gamma + \alpha B + \beta A + 2AB}{z(z-1)} + \frac{\delta + \alpha A + A(A-1)}{z^2} + \frac{\epsilon + \beta B + B(B-1)}{(z-1)^2}. \tag{9.2.17}$$

Choosing A and B as roots of $A(A-1) + \alpha A + \delta = 0$ and $B(B-1) + \beta B + \epsilon = 0$ in these equations then gives:

$$P(z) = \frac{a'}{z} + \frac{b'}{z-1} = \frac{a' - (a'+b')z}{z(1-z)} \tag{9.2.18}$$

and

$$Q(z) = \frac{c'}{z(1-z)}. \tag{9.2.19}$$

Using these functions in the O.D.E. (9.2.11) shows that $v(z)$ satisfies eq. (9.2.6) as was to be proven. The three arbitrary constants a', b' and c' may be relabelled in terms of the conventional constants a, b, c appearing in eq. (9.2.6). The required redefinitions are: $a' = c$, $b' = 1 + a + b - c$ and $c' = -ab$. □

We could perservere in this fashion indefinitely and spell out the most general equation having any given number of regular singular points. With the folk theorem of the introduction in mind, however, we do not do so since all of the equations of interest for our purposes are included in the Hypergeometric equation or its confluent counterpart to which we now turn.

9.3 The Confluent Hypergeometric Equation

According to the folk theorem, the only other type of differential equations that need be considered are those with one regular singular point and one irregular singular point. As usual the freedom to change variables may always be used to place the regular singular point at $z = 0$ and the irregular one at $z = \infty$. Unfortunately there is no analogue of the results of the previous section for this type of equation. It is therefore fortunate that we need consider only those equations obtained from an O.D.E. with three regular singular points by coalescing two of the regular points together into an irregular one. Since we have the most general such equation with three regular singular points, the general form of the resulting *confluent* equation may also be constructed.

To do so, start with the hypergeometric equation with regular singularities at $z = 0$, $z = 1$ and $z = \infty$. The point at $z = 1$ may be moved out to infinity while keeping the other two singular points fixed *via* the transformation $w = z/\lambda$ with the limit $\lambda \to 0$ taken at the end. Performing this change of variables on the Hypergeometric equation, (9.2.6), gives:

$$w(1 - \lambda w)\frac{d^2y}{dw^2} + [c - (1 + a + b)\lambda w]\frac{dy}{dw} - ab\lambda y(w) = 0. \tag{9.3.1}$$

An overall factor of $1/\lambda$ has been dropped in this equation to ensure that the second-derivative term has a finite limit as $\lambda \to 0$ with both w and y fixed. We need to know how the parameters a, b and c should vary in passing to this limit. Our guiding principle is to try to obtain the least restrictive equation possible. The least restrictive choice is easily seen to be to take $\lambda \to 0$ with a and c fixed but with $b \to \infty$ such that $b\lambda$ is fixed. (Notice that eq. (9.3.1) is invariant under switching a with b so it is immaterial whether it is a or b that is taken to infinity.) Since the fixed value of $b\lambda$ may be absorbed by rescaling the independent variable w there is no loss of generality in choosing $b = 1/\lambda$.

The equation obtained in this limit is the *Confluent Hypergeometric Equation*:

$$zy''(z) + (c - z)y'(z) - ay(z) = 0. \tag{9.3.2}$$

9.4 Connection to Commonly Occuring Equations

The content of the folk theorem is that for most purposes it suffices to study the two linear, second-order O.D.E.'s (9.2.5) and (9.3.2). All of the commonly occuring O.D.E.'s of mathematical physics correspond to special cases of these two equations with specific values for the parameters a, b and c. The purpose of the present section is to make this connection explicit for the equations that have been encountered so far: Bessel's equation and the Associated Legendre equation. These are chosen both because they occur most frequently in physical problems and because they are representative examples of the techniques used in making this connection.

EXAMPLE 1: Consider first the Associated Legendre equation, (5.3.10):

$$y''(z) - \frac{2z}{1-z^2}y'(z) - \left[\frac{t}{1-z^2} + \frac{m^2}{(1-z^2)^2}\right]y(z) = 0 \tag{9.4.1}$$

m is an integer and t is a complex number.

This has exactly three regular singular points, located at $z = \pm 1$ and $z = \infty$, and so falls into category III above. It must therefore be a special case of the Hypergeometric equation. In order to establish this connection and to determine the corresponding values for a, b and c eq. (9.4.1) must be put into the standard form (9.2.6). This means first putting the singular points at 0,1 and ∞ and then changing the dependent variable to put $q(z)$ into standard form.

The first part is simple: $w = \frac{1}{2}(1-z)$ is a fractional-linear transformation that puts the singularities at $w = 0, 1$ and ∞. The resulting form for the coefficient functions $p(w)$ and $q(w)$ is:

$$p(w) = \frac{1-2w}{w(1-w)} = \frac{1}{w} - \frac{1}{1-w} \tag{9.4.2}$$

and

$$q(w) = -\frac{t}{w(1-w)} - \frac{\frac{1}{4}m^2}{w^2(1-w)^2} = -\frac{t+\frac{1}{2}m^2}{w(1-w)} - \frac{\frac{1}{4}m^2}{w^2} - \frac{\frac{1}{4}m^2}{(1-w)^2}. \tag{9.4.3}$$

Defining the new dependent variable by $y(w) = w^A(w-1)^B v(w)$ changes these coefficient functions to:

$$P(w) = \frac{1+2A}{w} - \frac{1+2B}{1-w} \tag{9.4.4}$$

and

$$Q(w) = -\frac{t+\frac{1}{2}m^2+A+B+2AB}{w(1-w)} + \frac{A^2-\frac{1}{4}m^2}{w^2} + \frac{B^2-\frac{1}{4}m^2}{(1-w)^2}. \tag{9.4.5}$$

The choice $A^2 = B^2 = \frac{1}{4}m^2$ puts the O.D.E. into Hypergeometric form. Choose, then, $A = B = \frac{1}{2}m$. (The alternative choice $A = -B = \frac{1}{2}m$ gives a different but equally good representation of (9.4.1) as a Hypergeometric equation.) Then (9.4.4) and (9.4.5) become:

$$P(w) = \frac{1+m}{w} - \frac{1+m}{1-w} = \frac{(1+m)(1-2w)}{w(1-w)} \tag{9.4.6}$$

and

$$Q(w) = -\frac{t+m+m^2}{w(1-w)}. \tag{9.4.7}$$

The O.D.E. satisfied by $v(w)$ is then:

$$w(1-w)v''(w) + (1+m)(1-2w)v'(w) - (t+m+m^2)v(w) = 0 \tag{9.4.8}$$

which is in Hypergeometric form with $a+b = 1+2m$, $ab = t+m+m^2$ and $c = 1+m$. This has solutions $a = \frac{1}{2}(1+2m \pm \sqrt{1-4t})$ and $b = \frac{1}{2}(1+2m \mp \sqrt{1-4t})$. It does not matter whether the upper or lower sign is chosen in these expressions for a and b since changing this sign just has the effect of switching a and b and the Hypergeometric equation is unchanged by such a switch.

The relation between the solutions to this O.D.E. and those of the original O.D.E. is:

$$y = [w(w-1)]^{m/2}v(w) = 2^{-m}e^{i\pi m/2}(1-z^2)^{m/2}v\left[\frac{1}{2}(1-z)\right]. \tag{9.4.9}$$

EXAMPLE 2: Bessel's equation, (5.2.9):

$$y''(z) + \frac{1}{z}y'(z) + \left(1 - \frac{\nu^2}{z^2}\right)y(z) = 0. \tag{9.4.10}$$

ν is in general a complex constant.

This equation has two singular points, a regular one at $z = 0$ and an irregular one at $z = \infty$. It is therefore potentially equivalent to the Confluent Hypergeometric equation (9.3.2). In principle, the strategy to be followed is the same as for the Hypergeometric equation. We perform the most general change of variables $u = u(z)$ and $y(u) = f(u)v(u)$ consistent with the requirement that (i) no new singularities are introduced, and (ii) that the existing singularities not be moved from their standard places ($z = 0$ and $z = \infty$ in the present example). The requirement that the O.D.E. in the new variables have the standard form then imposes conditions on the unknown functions $f(z)$ and $u(z)$ that can be solved. The resulting values for the constants a and c in the standard form can then be read off by inspection.

Unfortunately, since there is no general theorem stating that any O.D.E. with one regular and one irregular singular point can be put into Confluent Hypergeometric form, it may be that there is no change of variables that can transform the O.D.E. to the desired form. If not, then the complete set of conditions required of $u(z)$ and $f(z)$ will be found to have no solution.

In the present example we are allowed to perform any one-to-one transformation $u(z)$ of the complex plane to itself that does not move the singular points: $u(0) = 0$, $u(\infty) = \infty$. Furthermore, $u(z)$ must be analytic everywhere, except potentially at $z = 0$ and $z = \infty$, if it is not to introduce any new singular points apart from those at $z = 0$ and $z = \infty$. At $z = 0$ the potentially singular behaviour of $u(z)$ is constrained by the requirement that the singular point not become irregular once expressed in terms of the new variable.

In order to do this analysis properly an analogue of the theorems (1) and (2) of section 9.2 are required. These are more difficult to identify in this case since the behaviour as $z \to \infty$

is not restricted in the present example. Rather than trying to derive the most general transformation, then, we content ourselves with an educated guess of which transformations to try. To this end, notice that one type of transformation that satisfies all of the above conditions, though not necessarily the most general type, is furnished by the fractional-linear transformations $u = (Az + B)/(Cz + D)$. The requirement that the points $z = 0$ and $z = \infty$ not be moved then implies that $B = C = 0$ so the transformation becomes $u = Kz$ where K is an arbitrary complex constant.

Together with the change of dependent variable, $y(u) = f(u)v(u)$, the O.D.E. (9.4.10) becomes:

$$v''(u) + P(u)v'(u) + Q(u)v(u) = 0 \tag{9.4.11}$$

in which the coefficient functions are given by:

$$P(u) = \frac{1}{u} + 2\frac{f'}{f}(u) \tag{9.4.12}$$

and

$$Q(u) = \frac{1}{K^2} - \frac{\nu^2}{u^2} + \frac{1}{u}\frac{f'}{f}(u) + \frac{f''}{f}(u). \tag{9.4.13}$$

$f(u)$ and K are now to be determined by the requirement that $P(u) = (c/u) - 1$ and $Q(u) = -a/u$, since these are the coefficient functions for the standard Confluent Hypergeometric equation. A differential equation for $f(u)$ is obtained by equating eq. (9.4.12) to $(c/u) - 1$:

$$2\frac{f'}{f}(u) = \frac{c-1}{u} - 1. \tag{9.4.14}$$

This is easily integrated to yield (up to an irrelevant multiplicative constant):

$$f(u) = u^{(c-1)/2}e^{-u/2}. \tag{9.4.15}$$

Inserting this into eq. (9.4.13) for $Q(u)$ and equating the result to $-a/u$ gives:

$$\left[\frac{1}{4} + \frac{1}{K^2}\right] + \left[\frac{1-c}{2} - \frac{1}{2} + a\right]\frac{1}{u} + \left[\frac{(1-c)^2}{4} - \nu^2\right]\frac{1}{u^2} = 0. \tag{9.4.16}$$

The vanishing of the coefficient of each power of $1/u$ in this equation furnishes three equations that are to be solved for the three unknowns a, c and K. The solutions are: $K = 2i$, $c = 1+2\nu$ and $a = \frac{1}{2}c = \frac{1}{2} + \nu$. (Notice that, since Bessel's equation is unchanged by $\nu \to -\nu$, an alternate solution is given by $K = 2i$, $c = 1 - 2\nu$ and $a = \frac{1}{2} - \nu$.)

The conclusion, therefore, is that the change of variables

$$y(z) = z^{\pm\nu}e^{-iz}v(2iz) \tag{9.4.17}$$

reduces Bessel's equation to Confluent Hypergeometric form with parameters given by $c = 2a = 1 \pm 2\nu$.

10 Special Functions

10.1 Introduction

The classification theorem, together with the folk theorem concerning the type of O.D.E.'s that arise in mathematical physics, implies that once the solution to the standard form differential equations (9.2.6) and (9.3.2) are known, the solutions appropriate to most physical problems may be obtained as special cases. The present chapter is therefore devoted to applying the techniques of chapter 8 to solve these two master equations. The solutions so generated are, of course, in series form and in general cannot be expressed in closed form in terms of elementary functions. As a result much of the chapter is devoted to constructing analytic continuations of the solutions to the general complex plane and using these to derive the properties that are useful in manipulating these functions in physical problems. The general properties are then written explicitly for the special cases of Bessel's and Legendre's equations since these are the most commonly encountered and furnish classic examples of the techniques described.

10.2 Hypergeometric Functions

The aim is to explicitly generate a series solution to the Hypergeometric equation (9.2.6):

$$z(1-z)y''(z) + [c - (1+a+b)z]y'(z) - aby(z) = 0. \tag{10.2.1}$$

Since the coefficient functions have a simple form in powers of z it is convenient to expand about the regular singular point at $z = 0$. From Fuchs's theorem it follows that the solution must have the following form:

$$y(z) = \sum_{n=0}^{\infty} y_n z^{n+s} \tag{10.2.2}$$

which, when substituted into the differential equation gives the following indicial equation:

$$I(s) = s(s-1+c) = 0 \tag{10.2.3}$$

and recursion relation:

$$y_{n+1} = \left[\frac{(n+s+a)(n+s+b)}{(n+s+1)(n+s+c)}\right] y_n \qquad n \geq 0. \tag{10.2.4}$$

The roots of the indicial equation are $s_1 = 0$ and $s_2 = 1 - c$. From the general discussion in section 8.5 the series ansatz may fail to generate both solutions if c should be an integer. Recalling the property $\Gamma(z+1) = z\Gamma(z)$ satisfied by Euler's gamma function allows the general solution to the recursion relation (10.2.4) to be written down immediately:

$$y_n(s) = C\left[\frac{\Gamma(n+s+a)\Gamma(n+s+b)}{\Gamma(n+s+1)\Gamma(n+s+c)}\right]. \tag{10.2.5}$$

The constant C is determined to be $y_0[\Gamma(s+1)\Gamma(s+c)]/[\Gamma(s+a)\Gamma(s+b)]$ by the initial condition that eq. (10.2.5) reduce to the free parameter y_0 when $n=0$. Eq. (10.2.5) then reduces to:

$$y_n(s) = \left[\frac{\Gamma(c+s)\Gamma(s+1)}{\Gamma(a+s)\Gamma(b+s)}\right]\left[\frac{\Gamma(n+s+a)\Gamma(n+s+b)}{\Gamma(n+s+c)\Gamma(n+s+1)}\right] y_0. \tag{10.2.6}$$

Inspection of the limit

$$\lim_{n\to\infty} \frac{y_{n+1}(s) z^{s+n+1}}{y_n(s) z^{s+n}} = \lim_{n\to\infty}\left[\frac{(n+s+a)(n+s+b)}{(n+s+1)(n+s+c)}\right] z = z, \tag{10.2.7}$$

together with the ratio test, implies that for any a, b, c and s the series in eq. (10.2.2) converges for $|z|<1$.

The solution corresponding to $s=0$ is now easily written. With the conventional choice that $y_0=1$ the series solution using $s=0$ in eq. (10.2.6) is:

$$\begin{aligned} y_1(z) &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=0}^{\infty}\frac{\Gamma(n+a)\Gamma(n+b)}{\Gamma(n+c)n!}z^n \\ &= 1 + \frac{ab}{c}z + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{z^2}{2} + \dots \\ &\equiv F(a,b;c;z) \\ &\equiv {}_2F_1(a,b;c;z). \end{aligned} \tag{10.2.8}$$

The function $F(a,b;c;z)$ defined by this series is called the *Hypergeometric function* and its definition makes sense provided that $c \neq 0,-1,-2,\dots$ and $|z|<1$.

The second solution corresponds to the choice $s = 1-c$. The corresponding solution is therefore given by:

$$\begin{aligned} y_2(z) &= \frac{\Gamma(2-c)}{\Gamma(a+1-c)\Gamma(b+1-c)}\sum_{n=0}^{\infty}\frac{\Gamma(n+a+1-c)\Gamma(n+b+1-c)}{\Gamma(n+2-c)n!}z^{n+1-c} \\ &= z^{1-c}\left[1 + \frac{(a+1-c)(b+1-c)}{2-c}z + \dots\right] \\ &= z^{1-c}F(a+1-c, b+1-c; 2-c; z) \end{aligned} \tag{10.2.9}$$

This second solution is well defined only if $c \neq 2,3,\dots$ and $|z|<1$. If $c=1$ then the solutions (10.2.8) and (10.2.9) are not distinct. For non-integral c the two solutions $y_1(z)$ and $y_2(z)$ are linearly independent since they behave differently as $z \to 0$ and so cannot be proportional to one another.

10.3 Confluent Hypergeometric Functions

The series solution to the Confluent Hypergeometric equation (9.3.2):

$$zy''(z) + (c-z)y'(z) - ay(z) = 0 \tag{10.3.1}$$

can be obtained in a similar fashion. It is more instructive, however, to obtain it directly from the previously constructed Hypergeometric function. To do so take $u = z/\lambda$ and $b = 1/\lambda$ in eq. (10.2.6). The solution to the Confluent equation is obtained by taking the limit as $\lambda \to 0$ with all other quantities fixed:

$$\begin{aligned} y_1(u) &\equiv M(a, c; u) \\ &\equiv {}_1F_1(a; c; u) \\ &= \lim_{\lambda \to 0} F(a, 1/\lambda; c; \lambda u) \end{aligned} \tag{10.3.2}$$

for $c \neq 0, -1, ...$ The required limit for the $n+1$'th term of the series is:

$$\begin{aligned} X &\equiv \lim_{\lambda \to 0} \left[\lambda^n \frac{\Gamma(n + 1/\lambda)}{\Gamma(1/\lambda)} \right] \\ &= \lim_{\lambda \to 0} \left[\lambda^n \left(\frac{1}{\lambda} \right) \left(\frac{1}{\lambda} + 1 \right) ... \left(\frac{1}{\lambda} + n - 1 \right) \right] \\ &= \lim_{\lambda \to 0} [1(1 + \lambda)(1 + 2\lambda)...(1 + (n-1)\lambda)] \\ &= 1. \end{aligned} \tag{10.3.3}$$

This leaves the following series solution to eq. (10.3.1):

$$\begin{aligned} M(a, c; z) &= \left[\frac{\Gamma(c)}{\Gamma(a)} \right] \sum_{n=0}^{\infty} \left[\frac{\Gamma(a + n)}{\Gamma(c + n)} \right] \frac{z^n}{n!} \\ &= 1 + \left[\frac{a}{c} \right] z + \left[\frac{a(a+1)}{c(c+1)} \right] \frac{z^2}{2} + ... \end{aligned} \tag{10.3.4}$$

for $c \neq 0, -1, ...$ This function is called the *Confluent Hypergeometric function.*

Using the limit:

$$\lim_{n \to \infty} \frac{y_{n+1} z^{n+1}}{y_n z^n} = \lim_{n \to \infty} \left[\frac{(a+n)}{(c+n)(n+1)} \right] z = 0 \tag{10.3.5}$$

in the ratio test implies that the series (10.3.4) has an infinite radius of convergence.

The second solution is similarly found to be given by:

$$y_2(z) = z^{1-c} M(a + 1 - c, 2 - c; z) \tag{10.3.6}$$

provided $c \neq 2, 3, 4....$ For $c = 1$ solutions (10.3.2) and (10.3.6) are not distinct. For noninteger c these solutions are well-defined for all finite z and linearly independent.

EXAMPLES:

From the results of chapter 9 we know that the solutions to Bessel's equation and the Associated Legendre equation can be directly expressed in terms of the Hypergeometric and Confluent Hypergeometric series. Eq. (9.4.9) relates the solutions for the Associated Legendre equation:

$$y_1(z) = C_m \left(\frac{1-z}{1+z} \right)^{m/2} F \left[\frac{1}{2}(1 + \sqrt{1-4t}), \frac{1}{2}(1 - \sqrt{1-4t}); 1 + m; \frac{1}{2}(1-z) \right]. \tag{10.3.7}$$

with $m \neq -1, -2, ...$, to the Hypergeometric function. C_m denotes a constant. If $m < 0$ the solution can be taken as eq. (10.3.7) with m replaced everywhere by $-m$ since the Associated Legendre equation is invariant under this substitution. This is equivalent to the expression (10.2.9) for $y_2(z)$. Since the two roots to the indicial equation differ by $c = 1 + m$, the second linearly independent solution does not have the simple series form and must be constructed using the techniques of section 8.5.

Similarly, referring to eq. (9.4.17) shows that the solutions to Bessel's equation are given in terms of the confluent hypergeometric series by:

$$y_1(z) = C_\nu z^\nu e^{-iz} M \left[\frac{1}{2} + \nu, 1 + 2\nu; 2iz\right] \tag{10.3.8}$$

for $\nu \neq -\frac{1}{2}, -1, -\frac{3}{2}, ...$ C_ν again denotes a constant. The second solution is simply found by taking $\nu \to -\nu$ in this equation:

$$y_2(z) = C_{-\nu} z^{-\nu} e^{-iz} M \left[\frac{1}{2} - \nu, 1 - 2\nu; 2iz\right] \tag{10.3.9}$$

provided $\nu \neq \frac{1}{2}, 1, \frac{3}{2}, ...$

This construction is not limited to the Legendre or Bessel equations. To illustrate this point and for convenience of reference, the solution to the most frequently occuring O.D.E.'s of mathematical physics are briefly listed here in terms of Hypergeometric or Confluent Hypergeometric series:

1. ULTRASPHERICAL (GEGENBAUER) EQUATION:

$$(1 - z^2)y''(z) - (2\beta + 1)zy'(z) + n(n + 2\beta)y(z) = 0. \tag{10.3.10}$$

The regular solutions to this equation are denoted ${T_n}^\beta(z)$ and are called *Ultraspherical functions*. They are related to the Hypergeometric series by:

$${T_n}^\beta(z) = \frac{\Gamma(n + 2\beta + 1)}{2^\beta n!\Gamma(\beta + 1)} F \left[-n, n + 2\beta + 1; 1 + \beta; \frac{1}{2}(1 - z)\right]. \tag{10.3.11}$$

2. ASSOCIATED LEGENDRE FUNCTIONS:

$$y''(z) - \frac{2z}{1 - z^2}y'(z) + \left[\frac{\ell(\ell + 1)}{1 - z^2} - \frac{m^2}{(1 - z^2)^2}\right] y(z) = 0. \tag{10.3.12}$$

The regular solutions are the *Associated Legendre functions*, denoted ${P_\ell}^m(z)$:

$${P_\ell}^m(z) = \frac{1}{2^m m!} \frac{(\ell + m)!}{(\ell - m)!}(1 - z^2)^{m/2} F \left[m - \ell, m + \ell + 1; m + 1; \frac{1}{2}(1 - z)\right]. \tag{10.3.13}$$

This result is derived in sections 9.4 and 10.6. Notice that these are special cases of ultraspherical functions.

3. CHEBYSHEV POLYNOMIALS:

$$y''(z) - \frac{z}{1-z^2}y'(z) + \frac{n^2}{1-z^2}y(z) = 0. \tag{10.3.14}$$

has regular solutions, $T_n(z)$, called *Chebyshev polynomials*:

$$T_n(z) = F\left[-n, n; \frac{1}{2}; \frac{1}{2}(1-z)\right]. \tag{10.3.15}$$

These are again special cases of the ultraspherical functions.

4. BESSEL FUNCTIONS:

$$y''(z) + \frac{1}{z}y'(z) + \left(1 - \frac{\nu^2}{z^2}\right)y(z) = 0. \tag{10.3.16}$$

has as regular solutions the *Bessel functions* $J_\nu(z)$. As shown in sections 9.4 and 10.7 these are given by:

$$J_\nu(z) = \frac{1}{\Gamma(\nu+1)}\left(\frac{z}{2}\right)^\nu e^{-iz} M\left[\nu + \frac{1}{2}, 2\nu+1; 2iz\right]. \tag{10.3.17}$$

5. HERMITE FUNCTIONS:

$$y''(z) - 2zy'(z) + 2ny(z) = 0. \tag{10.3.18}$$

has *Hermite functions*, $H_n(z)$, as regular solutions:

$$H_{2n}(z) = (-)^n \frac{(2n)!}{n!} M\left[-n, \frac{1}{2}; z^2\right] \tag{10.3.19}$$

$$H_{2n+1}(z) = (-)^n \frac{2(2n+1)!}{n!} zM\left[-n, \frac{3}{2}; z^2\right]. \tag{10.3.20}$$

6. ASSOCIATED LAGUERRE FUNCTIONS:

$$y''(z) + \frac{k+1-z}{z}y'(z) + \frac{n}{z}y(z) = 0. \tag{10.3.21}$$

The solutions are the *Associated Laguerre functions*, ${L_n}^k(z)$, given by:

$${L_n}^k(z) = \frac{(n+k)!}{n!k!} M[-n, k+1; z]. \tag{10.3.22}$$

Having established a series representation for the solutions to the master O.D.E.'s and having shown how to relate the general solutions to the various differential equations generally encountered, it is necessary to develop some of the properties common to all of these functions. These can then be specialized to the specific cases of interest. The following two sections are therefore devoted to developing some general properties of the Hypergeometric and Confluent Hypergeometric functions. The purpose of this presentation is to (i) derive alternate representations of these functions for later use and (ii) illustrate the general procedure through which the properties of all such functions are derived.

The detailed application of these properties to Legendre and Bessel functions is then taken up in sections 10.6 and 10.7 respectively.

10.4 Integral Representations

This section is devoted to the development of a very useful set of integral representations for the Hypergeometric and Confluent Hypergeometric functions. The utility of these representations lies in their use as a method for analytically continuing the series form of these functions as well as for their central role in the demonstration of other properties, such as asymptotic forms *etc.*.

Consider first the Hypergeometric function, $F(a,b;c;z)$.

Theorem 10.1. *An integral representation for $F(a,b;c;z)$ is given by:*

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 dt\; t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}. \tag{10.4.1}$$

This converges for all z not lying between 1 and ∞ on the positive real axis if $Re\, c > Re\, b > 0$.

Proof. It suffices to show that the integral (10.4.1) agrees with the series definition (10.2.8) within its disc of convergence $|z| < 1$. Once this is done the integral representation furnishes an analytic continuation of $F(a,b;c;z)$ out of the unit disc. To establish the result expand the factor $(1-tz)^{-a}$ of the integrand in powers of z. For $t \leq 1$ the resulting binomial series:

$$\begin{aligned}(1-tz)^{-a} &= \sum_{n=0}^{\infty} \frac{\Gamma(1-a)}{\Gamma(1-a-n)n!}(-tz)^n \\ &= \sum_{n=0}^{\infty} \frac{\Gamma(a+n)}{\Gamma(a)n!}(tz)^n \end{aligned} \tag{10.4.2}$$

converges uniformly for $|z| < 1$ and so may be integrated term-by-term everywhere within the unit disc. To establish the second equality in eq. (10.4.2) use the following argument:

$$\begin{aligned}\frac{\Gamma(1-a)}{\Gamma(1-a-n)} &= (-a)(-a-1)...(-a-n+1) \\ &= (-)^n a(a+1)...(a+n-1) = (-)^n \frac{\Gamma(a+n)}{\Gamma(a)} \end{aligned} \tag{10.4.3}$$

Inserting (10.4.2) into the right-hand-side of (10.4.1) gives :

$$\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(c-b)}\sum_{n=0}^{\infty}\frac{\Gamma(a+n)}{n!}z^n\int_0^1 dt\; t^{n+b-1}(1-t)^{c-b-1}. \tag{10.4.4}$$

The integral may be evaluated and is equal to Euler's beta function $\Gamma(n+b)\Gamma(c-b)/\Gamma(n+c)$. Inserting this into eq. (10.4.4) gives:

$$\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=0}^{\infty}\frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)n!}z^n \tag{10.4.5}$$

which may be recognized as the series representation of $F(a,b;c;z)$ as claimed. $\square$

The integral representation (10.4.1) extends the definition of $F(a,b;c;z)$ but is still subject to the restrictions Re $c >$ Re $b > 0$ and z not real and greater than one. These restrictions come from the conditions for the convergence of the integral at the singular points of the integrand: $t = 0$, $t = 1$ and $t = 1/z$. For this reason it is worth rewriting the integral representation in a fashion that allows these conditions to be relaxed. In preparation for doing so change variables in the integral (10.4.1) to $s = 1/t$. The integral representation becomes:

$$F(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_1^{\infty} ds\; s^{a-c}(s-1)^{c-b-1}(s-z)^{-a}. \tag{10.4.6}$$

The idea is to now consider the same integrand as in eq. (10.4.6) but to change the contour of integration so that it does not pass through the singular points of the integrand: $s = 0$, $s = 1$ and $s = z$. That is consider the integral:

$$\mathcal{I}(z) = \int_{\gamma} s^{a-c}(s-1)^{c-b-1}(s-z)^{-a}\; ds \tag{10.4.7}$$

in which the integration contour γ is some curve in the complex s-plane not passing through 0,1 or z. The hope is that, subject to certain still-to-be-specified conditions on γ, this integral still furnishes an integral representation of $F(a,b;c;z)$ but for more general values of b, c and z since the potentially divergent points have been excluded from the integration path.

To determine whether $\mathcal{I}(z)$ is a representation of the Hypergeometric function operate on it with the Hypergeometric differential operator:

$$\mathcal{L} = z(1-z)\frac{d^2}{dz^2} + [c-(1+a+b)z]\frac{d}{dz} - ab \tag{10.4.8}$$

and see if the result vanishes. Directly differentiating (10.4.7) under the integral sign gives:

$$\mathcal{L}\mathcal{I}(z) = a\int_{\gamma} ds\;\frac{d}{ds}\left[s^{a-c+1}(s-1)^{c-b}(s-z)^{-a-1}\right]. \tag{10.4.9}$$

This is evidently zero if and only if the integral of the total derivative on the right-hand-side is zero. To see what this means in more detail, consider the original contour: $\gamma_o = (1,\infty)$. In this case eq. (10.4.9) reduces to:

$$\mathcal{L}\mathcal{I}(z) = a\left[s^{a-c+1}(s-1)^{c-b}(s-z)^{-a-1}\right]_{s=1}^{s=\infty}. \tag{10.4.10}$$

The contribution to these endpoint terms from $s = 1$ vanishes if and only if Re $c >$ Re b. Similarly, the contribution from $s = \infty$ is zero if Re $b > 0$. These are precisely the conditions demanded earlier for convergence of the integral.

These considerations may be summarized by the following

Theorem 10.2. *$\mathcal{I}(z)$ as defined by eq. (10.4.7) furnishes an integral representation for the Hypergeometric function $F(a, b; c; z)$:*

$$F(a, b; c; z) = C\mathcal{I}(z) \tag{10.4.11}$$

for some constant C provided that the following two properties are satisfied by a, b, c and γ:

1. *The right-hand-side of eq. (10.4.9) vanishes for all z.*
2. *The integral (10.4.7) converges, is not identically zero, and is analytic at $z = 0$.*

Proof. The proof is straightforward. If the integral (10.4.7) converges and is not identically zero then, by (1), it must be a nontrivial linear combination of the two basis solutions, $y_1(z)$ and $y_2(z)$, to the Hypergeometric equation defined by eqs. (10.2.8) and (10.2.9) respectively. Since $y_2(z)$ behaves like z^{1-c} as $z \to 0$ while $y_1(z)$ is analytic there, the condition that $\mathcal{I}(z)$ be analytic at $z = 0$ ensures that it is proportional to $y_1(z) = F(a, b; c; z)$, as required. The value of the constant C is determined to be $C = [\mathcal{I}(z = 0)]^{-1}$. □

The utility of this alternate form for the integral representation is that it is often possible to choose γ to satisfy conditions (1) and (2) without requiring conditions on a, b, c or z. For example, if the contour γ is a closed loop, then the vanishing of the surface term follows automatically. If the contour is chosen to avoid the three singular points of the integrand then it also generally converges and defines an analytic function of z. All that is required is to choose the contour such that the integral does not vanish identically. Just such a contour will be of interest for the special case of the Legendre functions in section 10.6.

Analogous expression may be derived for the Confluent Hypergeometric functions. An analogue of the representation (10.4.1) may, as usual, be found by replacing $z \to \lambda z$ and $b \to 1/\lambda$ followed by the limit $\lambda \to 0$. In equation (10.4.1) it is convenient to first relabel $a \leftrightarrow b$ before performing this substitution. The required limit is:

$$\lim_{\lambda \to 0} (1 - \lambda x)^{-\frac{1}{\lambda}} = e^x,$$

giving the following

Theorem 10.3.

$$M(a, c; z) = \frac{\Gamma(c)}{\Gamma(c-a)\Gamma(a)} \int_0^1 dt\, e^{zt} t^{a-1} (1-t)^{c-a-1} \tag{10.4.12}$$

This integral converges for all z provided that Re $c >$ Re $a > 0$.

Proof. The proof proceeds just as for the Hypergeometric case. The Taylor series for the exponential inside the integral is $\sum(zt)^n/n!$ and converges throughout the z-plane. Inserting this into eq. (10.4.12) and performing the integral term-by-term gives the series representation eq. (10.3.4). $\square$

Since the original series definition, eq. (10.3.4), of $M(a,c;z)$ converges for all z, a and c the integral representation eq. (10.4.12) does not extend the domain of definition of this function. It is nevertheless of interest because it is more useful than the series form for some purposes. As an example, the large-z behaviour of $M(a,c;z)$ is easily obtained from the representation (10.4.12) (but not from the series form!) by evaluating the integral by steepest descent.

The generalization of this representation to a more general contour, γ, follows immediately.

Theorem 10.4. *$M(a,c;z)$ admits the integral representation:*

$$M(a,c;z) = C \int_\gamma dt\; e^{zt} t^{a-1} (1-t)^{c-a-1} \tag{10.4.13}$$

provided that:

1. *The integral converges, does not vanish identically and is analytic at $z = 0$.*
2. *The following quantity vanishes:*

$$\int_\gamma dt\; \frac{d}{dt}\left[t^a (1-t)^{c-a} e^{zt}\right]. \tag{10.4.14}$$

Proof. The proof is as before. The integral (10.4.14) is the result of applying the confluent Hypergeometric differential operator to the integral and so condition (2) is the requirement that the integral satisfy the Confluent Hypergeometric equation. Condition (1) ensures that the solution is proportional to the specific solution $M(a,c;z)$. The constant C is determined by requiring that the result equal unity when $z = 0$. $\square$

10.5 Recurrence Relations

The solutions $F(a,b;c;z)$ and $M(a,c;z)$ satisfy a wide class of identities that relate functions (and sometimes their derivatives) that have arguments a, b and c that differ by integers. These relations are often extremely useful when manipulating quantities such as integrals involving special functions. Two typical such relations are derived in this section to illustrate the general procedure for their generation.

Theorem 10.5. *The following identity holds for all a, b, c and z:*

$$M(a,c-1;z) = \frac{z}{c-1} M(a,c;z) + M(a-1,c-1;z). \tag{10.5.1}$$

Proof. The proof starts with the trivial identity:

$$0 = \int_0^1 dt\, \frac{d}{dt}\left[e^{zt}t^{a-1}(1-t)^{c-a-1}\right] \tag{10.5.2}$$

for Re $c >$ Re $a > 1$. The idea is to first prove the result for Re $c >$ Re $a > 1$ and then to infer its validity for other values of a an c by analytic continuation. To proceed, explicitly perform the differentiation and compare each resulting term to the integral representation (10.4.12):

$$\begin{aligned} 0 &= \int_0^1 dt\, \left[ze^{zt}t^{a-1}(1-t)^{c-a-1} + (a-1)e^{zt}t^{a-2}(1-t)^{c-a-1}\right. \\ &\qquad\qquad \left. -(c-a-1)e^{zt}t^{a-1}(1-t)^{c-a-2}\right] \\ &= z\frac{\Gamma(a)\Gamma(c-a)}{\Gamma(c)}M(a,c;z) + (a-1)\frac{\Gamma(a-1)\Gamma(c-a)}{\Gamma(c-1)}M(a-1,c-1;z) \\ &\qquad -(c-a-1)\frac{\Gamma(a)\Gamma(c-a-1)}{\Gamma(c-1)}M(a,c-1;z) \end{aligned} \tag{10.5.3}$$

Cancelling an overall factor of $\Gamma(a)\Gamma(c-a)/\Gamma(c)$ then gives (using $\Gamma(z+1) = z\Gamma(z)$) the desired result. □

The second representative identity (that proves useful in the following section devoted to Legendre functions) is:

Theorem 10.6.

$$\frac{d}{dz}F(a,b;c;z) = \frac{ab}{c}F(a+1,b+1;c+1;z). \tag{10.5.4}$$

Proof. This is proven by directly differentiating the defining series (10.2.8) for $F(a,b;c;z)$:

$$\begin{aligned} \frac{d}{dz}F(a,b;c;z) &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=1}^{\infty}\frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)}\frac{z^{n-1}}{(n-1)!} \\ &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=0}^{\infty}\frac{\Gamma(a+n+1)\Gamma(b+n+1)}{\Gamma(c+n+1)}\frac{z^n}{n!} \\ &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(c+1)}F(a+1,b+1;c+1;z) \\ &= \frac{ab}{c}F(a+1,b+1;c+1;z) \end{aligned} \tag{10.5.5}$$

□

Gauss' recursion relations

Many other identities can be derived in a similar way and have implications for all special functions that are related to hypergeometric functions. Here is a list (quoted without proof) of many such useful identities, called Gauss's recursion relations:

$$c[c-1-(2c-a-b-1)z]F(a,b;c;z)+(c-a)(c-b)zF(a,b;c+1;z) \\ +c(c-1)(z-1)F(a,b;c-1;z)=0\,, \tag{10.5.6}$$

$$(2a-c-az+bz)F(a,b;c;z)+(c-a)F(a-1,b;c;z)+a(z-1)F(a+1,b;c;z)=0\,, \tag{10.5.7}$$

$$(2b-c-bz+az)F(a,b;c;z)+(c-b)F(a,b-1;c;z)+b(z-1)F(a,b+1;c;z)=0\,, \tag{10.5.8}$$

$$cF(a,b-1;c;z)-cF(a-1,b;c;z)+(a-b)zF(a,b;c+1;z)=0\,, \tag{10.5.9}$$

$$c(a-b)F(a,b;c;z)-a(c-b)F(a+1,b;c+1;z)+b(c-a)F(a,b+1;c+1;z)=0\,, \tag{10.5.10}$$

$$c(c+1)F(a,b;c;z)-c(c+1)F(a,b;c+1;z)-abzF(a+1,b+1;c+2;z)=0\,, \tag{10.5.11}$$

$$cF(a,b;c;z)-(c-a)F(a,b+1;c+1;z)-a(1-z)F(a+1,b+1;c+1;z)=0\,, \tag{10.5.12}$$

$$cF(a,b;c;z)+(b-c)F(a+1,b;c+1;z)-b(1-z)F(a+1,b+1;c+1;z)=0\,, \tag{10.5.13}$$

$$c(c-bz-a)F(a,b;c;z)-c(c-a)F(a-1,b;c;z)+abz(1-z)F(a+1,b+1;c+1;z)=0\,, \tag{10.5.14}$$

$$c(c-az-b)F(a,b;c;z)-c(c-b)F(a,b-1;c;z)+abz(1-z)F(a+1,b+1;c+1;z)=0\,, \tag{10.5.15}$$

$$cF(a,b;c;z)-cF(a,b+1;c;z)+azF(a+1,b+1;c+1;z)=0\,, \tag{10.5.16}$$

$$cF(a,b;c;z)-cF(a+1,b;c;z)+bzF(a+1,b+1;c+1;z)=0\,, \tag{10.5.17}$$

$$c[a-(c-b)z]F(a,b;c;z)-ac(1-z)F(a+1,b;c;z)+(c-a)(c-b)zF(a,b;c+1;z)=0\,, \tag{10.5.18}$$

$$c[b-(c-a)z]F(a,b;c;z)-bc(1-z)F(a,b+1;c;z)+(c-a)(c-b)zF(a,b;c+1;z)=0\,, \tag{10.5.19}$$

$$c(c+1)F(a,b;c;z)-c(c+1)F(a,b+1;c+1;z)+a(c-b)zF(a+1,b+1;c+2;z)=0\,, \tag{10.5.20}$$

$$c(c+1)F(a,b;c;z)-c(c+1)F(a+1,b;c+1;z)+b(c-a)zF(a+1,b+1;c+2;z)=0\,, \tag{10.5.21}$$

$$cF(a,b;c;z)-(c-b)F(a,b;c+1;z)-bF(a,b+1;c+1;z)=0\,, \tag{10.5.22}$$

$$cF(a,b;c;z)-(c-a)F(a,b;c+1;z)-aF(a+1,b;c+1;z)=0\,, \tag{10.5.23}$$

and so on.

Hypergeometric group and transformation formulae

As was mentioned earlier, there is a class of six fractional-linear changes of variable that do not change the singularities or linearity of the Hypergeometric equation, but do swap the singular points. These are given in section 9.2, and are repeated here:

$$u = z\,, \quad u = \frac{1}{z}\,, \quad u = 1 - z\,, \quad u = \frac{1}{1-z}\,, \quad u = 1 - \frac{1}{z} \quad \text{and} \quad u = -\frac{z}{1-z}\,. \tag{10.5.24}$$

Because these transformations map the Hypergeometric equation onto itself they can be regarded as mapping the three parameters (a, b, c) into a new set (a', b', c'). For instance, substituting $u = 1 - z$ into the Hypergeometric equation

$$u(1-u)\,\frac{d^2y}{du^2} + \Big[c - (a+b+1)u\Big]\frac{dy}{du} - ab\,y = 0 \tag{10.5.25}$$

gives

$$z(1-z)\,\frac{d^2y}{dz^2} + \Big[(1+a+b-c) - (a+b+1)z\Big]\frac{dy}{dz} - ab\,y = 0 \tag{10.5.26}$$

which is also Hypergeometric, but with $a' = a$, $b' = b$ and $c' = 1 + a + b - c$.

The Hypergeometric function $F(a, b; c; z)$ is defined to be analytic at $z = 0$, but can be singular in the vicinity of the other two singular points $z = 1$ and $z \to \infty$. The Hypergeometric group provides a simple way to identify this singular behaviour because it maps solutions expanded about $z = 0$ onto solutions expanded about the other two singular points. Since any two linearly independent solutions provide a basis of solutions, the solutions expanded about the other singular points can always be written in terms of the solutions found by expanding about $z = 0$ and vice versa.

The resulting expressions are called transformation formulae, and can be efficiently found by changing variables in integral representations, like that given in (10.4.1):

$$F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 dt\; t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}. \tag{10.5.27}$$

For instance, performing the changes of integration variable $t \to 1 - t$, $t \to 1/(1 - z - tz)$ and $t \to (1 - t)/(1 - tz)$ in this expression leads to the following equivalent forms

$$\begin{aligned} F(a, b; c; z) &= (1-z)^{-a} F[a, c-b; c; z/(z-1)] \\ &= (1-z)^{-b} F[c-a, b; c; z/(z-1)] \\ &= (1-z)^{c-a-b} F(c-a, c-b; c; z)\,. \end{aligned} \tag{10.5.28}$$

Similar transformation formulae that can be obtained in this way are

$$\begin{aligned} F(a, b; c; z) = {}& \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}\, F(a, b; a+b-c+1; 1-z) \\ & + (1-z)^{c-a-b} \frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}\, F(c-a, c-b; c-a-b+1; 1-z)\,, \end{aligned} \tag{10.5.29}$$

which – using $F(a,b;c;0) = 1$ – displays the singularity structure of $F(a,b;c;z)$ near $z = 1$. Similarly,

$$F(a,b;\Gamma;z) = (1-z)^{-a}\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)}\,F[a,c-b;a-b+1;1/(1-z)] \tag{10.5.30}$$
$$+(1-z)^{-b}\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}\,F[b,c-a;b-a+1;1/(1-z)]\,,$$

or

$$F(a,b;c;z) = \frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)}\,(-z)^{-a}F[a,a+1-c;a+1-b;1/z] \tag{10.5.31}$$
$$+\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}(-z)^{-b}F[b,b+1-c;b+1-a;1/z]\,,$$

can be used to display the singularity structure as $z \to \infty$ (where this last expression assumes $|\arg z| < \pi$ and $a-b$ not an integer).

10.6 Legendre Functions

We turn to the exploration of the consequences of these properties in some detail for the Legendre functions. Recall that in section 9.4 it was shown that the solutions to the Associated Legendre equation:

$$\frac{d}{dz}\left[(1-z^2)\frac{dy}{dz}\right] - ty(z) - \frac{m^2}{1-z^2}y(z) = 0 \tag{10.6.1}$$

may be expressed in terms of the Hypergeometric functions by:

$$y = C(1-z^2)^{m/2}F\left[\frac{1}{2}(1+2m+\sqrt{1-4t}),\frac{1}{2}(1+2m-\sqrt{1-4t});1+m;\frac{1}{2}(1-z)\right] \tag{10.6.2}$$

for $m \neq -1,-2,\ldots.$ For negative m the solution may be taken to again be (10.6.2) but with m replaced by $-m$. Both of these cases may be considered at once in (10.6.2) if m is replaced everywhere by $|m|$. The second, linearly independent solution cannot be so expressed since $c = 1+m$ is an integer implying that the second solution involves a logarithmic singularity.

This solution is by construction bounded at $z = 1$, since both $F(a,b;c;0) = 1$ and $(1-z^2)^{|m|/2}$ are bounded there. The first property to be now established gives the circumstances under which this solution is also bounded at $z = -1$:

Theorem 10.7. *The only solutions to (10.6.1) that are bounded at both $z = 1$ and $z = -1$ are given by (10.6.2) with $t = -\ell(\ell+1)$ with ℓ being an integer that can take the potential values: $\ell = |m|, |m|+1, \ldots.$ If t does not satisfy this condition then there are no solutions that are bounded at both ± 1.*

Proof. In order to determine the behaviour of (10.6.2) at $z = -1$ the behaviour of the Hypergeometric function $F(a,b;c;u)$ at $u = \frac{1}{2}(1-z) = 1$ is required. To get this use the integral representation (10.4.1):

$$F(a,b;c;z=1) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 dt\; t^{b-1}(1-t)^{c-b-a-1} \tag{10.6.3}$$

which converges for Re $(c-a) >$ Re $b > 0$. Evaluating the integral in terms of Euler's beta function, eq.(6.5.3), gives:

$$\begin{aligned} F(a,b;c;z=1) &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \cdot \frac{\Gamma(b)\Gamma(c-b-a)}{\Gamma(c-a)} \\ &= \frac{\Gamma(c)\Gamma(c-b-a)}{\Gamma(c-b)\Gamma(c-a)} \end{aligned} \tag{10.6.4}$$

Now, the values $a = \frac{1}{2}(1+2|m|+\sqrt{1-4t})$, $b = \frac{1}{2}(1+2|m|-\sqrt{1-4t})$ and $c = 1+|m|$ appropriate to the Associated Legendre equation satisfy the relation: $c-a-b = -|m|$. Because of the factor $\Gamma(c-a-b)$ appearing in eq. (10.6.4) this implies that $F(a,b;c;z)$ always diverges for these values of a, b and c as $z \to 1$. The question becomes whether the prefactor $(1-z^2)^{|m|/2}$ vanishes quickly enough to overcome the divergence of $F(a,b;c;\frac{1}{2}(1-z))$ as $z \to -1$.

To answer this question the nature of the divergence as $u \to 1$ of $F(a,b;c;u)$ is required. This question can be answered independent of the detailed form of the solution. Recall that $u = 1$ is one of the regular singular points of the Hypergeometric O.D.E.. As such, it is possible to generate series solutions about $u = 1$ of the form $(u-1)^s \sum_{n=0}^{\infty} y_n (u-1)^n$. The power, s, of the prefactor governs how the basis solutions behave as $u \to 1$ and is determined by the roots of the indicial equation for this expansion: $I(s) = s(s-1) + p_0 s + q_0 = 0$. Here p_0 and q_0 are the coefficients of $(u-1)^{-1}$ in $p(u)$ and of $(u-1)^{-2}$ in $q(u)$ respectively. They may be directly read off from the O.D.E. (10.2.1) and are given by $p_0 = 1+a+b-c$ and $q_0 = 0$. For the Associated Legendre equation these become $p_0 = 1+|m|$ and $q_0 = 0$. The roots of the indicial equation for expansions about $u = 1$ are therefore $s = 0$ and $s = 1 - p_0 = -|m|$. This implies that any solution to the Hypergeometric equation that diverges as $u \to 1$ must do so like $(u-1)^{-|m|}$. It follows that the solutions to (10.6.1) must diverge like $(1+z)^{-|m|/2}$ as $z \to -1$.

The inescapable conclusion is therefore that the power series solutions to (10.6.1) for general values of t must diverge at one of $z = \pm 1$. The only way to evade this result would be for the series defining $F(a,b;c;u)$ to *terminate* since in this case the prefactor $(1-z^2)^{|m|/2}$ in eq. (10.6.2) always ensures a bounded result.

Question: for what values of t does the series for $F(a,b;c;u)$ terminate? To answer this recall the recursion relation (10.2.4) that defines the series:

$$y_{n+1} = \left[\frac{(n+a)(n+b)}{(n+c)(n+1)}\right] y_n, \qquad n \geq 0. \tag{10.6.5}$$

If this is to have solutions with $y_k \neq 0$ but $y_{k+1} = 0$, for k a non-negative integer, it follows that either a or b must equal $-k$. Using the known expressions for a and b in terms of t implies that:

$$\frac{1}{2}(1+2|m| \pm \sqrt{1-4t}) = -k \tag{10.6.6}$$

which can be solved for t giving:

$$t = \frac{1}{4}[1-(2\ell+1)^2] = -\ell(\ell+1) \tag{10.6.7}$$

where $\ell = |m| + k$ defines the non-negative integer ℓ. This establishes the result of the theorem. $\square$

In quantum mechanics the condition $t = -\ell(\ell+1)$ has the physical interpretation of quantization of angular momentum. The O.D.E. (10.6.1) becomes:

$$\frac{d}{dz}\left[(1-z^2)\frac{dy}{dz}\right] + \ell(\ell+1)y(z) - \frac{m^2}{1-z^2}y(z) = 0 \tag{10.6.8}$$

and the solutions are:

$$y(z) = C(1-z^2)^{|m|/2} F\left[|m|-\ell, 1+|m|+\ell; 1+|m|; \frac{1}{2}(1-z)\right]. \tag{10.6.9}$$

The standard notation for this solution is ${P_\ell}^m(z)$ once a conventional choice for the normalization constant, C, has been made. For $m \geq 0$:

$${P_\ell}^{-m}(z) \equiv (-)^m \frac{(\ell-m)!}{(\ell+m)!}{P_\ell}^m(z)$$

and

$${P_\ell}^m(z) \equiv \frac{1}{2^m m!}\frac{(\ell+m)!}{(\ell-m)!}(1-z^2)^{|m|/2} F\left[m-\ell, 1+m+\ell; 1+m; \frac{1}{2}(1-z)\right]. \tag{10.6.10}$$

This choice of normalization might appear odd, however it is motivated by property (10.6.13) below.

Definition 10.1. *The Legendre Polynomials, $P_\ell(z)$, are defined by:*

$$P_\ell(z) = {P_\ell}^0(z) = F\left[-\ell, 1+\ell; 1; \frac{1}{2}(1-z)\right]. \tag{10.6.11}$$

$F\left[-\ell, 1+\ell; 1; \frac{1}{2}(1-z)\right]$ is by construction a polynomial of degree ℓ in $(1-z)$ and so $P_\ell(z)$ is also a polynomial in z of degree ℓ. Inspection of the Hypergeometric series (10.2.8) gives the following explicit form for $P_\ell(z)$:

$$P_\ell(z) = \sum_{n=0}^{\ell}(-)^n \frac{(\ell+n)!}{(\ell-n)!(n!)^2}\left[\frac{1}{2}(1-z)\right]^n. \tag{10.6.12}$$

In particular $P_\ell(z=1) = 1$ for all ℓ.

Theorem 10.8. *For $m \geq 0$ the Associated Legendre functions are related to the Legendre polynomials by:*

$${P_\ell}^m(z) = (1-z^2)^{m/2}\frac{d^m}{dz^m}P_\ell(z). \tag{10.6.13}$$

Proof. The proof of this result is based on the recursion relation (10.5.4) proven in the previous section:

$$\frac{d}{dz}F(a,b;c;z) = \frac{ab}{c}F(a+1,b+1;c+1;z). \tag{10.5.4}$$

When applied to the Hypergeometric function appearing in eq. (10.6.10) this implies:

$$\begin{aligned} y(z) &\equiv F\left[m-\ell, 1+m+\ell; 1+m; \frac{1}{2}(1-z)\right] \\ &= \frac{(-2)m}{(m-\ell-1)(m+\ell)}\frac{d}{dz}F\left[m-\ell-1, m+\ell; m; \frac{1}{2}(1-z)\right] \\ &= \frac{(-2)^m m(m-1)...1}{(m-\ell-1)...(-\ell)(m+\ell)...(\ell+1)}\frac{d^m}{dz^m}F\left[-\ell, \ell+1; 1; \frac{1}{2}(1-z)\right] \\ &= \frac{2^m m!(\ell-m)!}{(\ell+m)!}\frac{d^m}{dz^m}F\left[-\ell, \ell+1; 1; \frac{1}{2}(1-z)\right] \end{aligned} \tag{10.6.14}$$

This last line, because of the choice made for the normalization constant C in eq. (10.6.10) above, gives the desired result when the definitions for $P_\ell(z)$ and $P_\ell{}^m(z)$ are inserted. □

This last result implies that the Associated Legendre functions may all be obtained by differentiating the Legendre polynomials. Their properties may be inferred once the corresponding properties of these polynomials are known. For this reason the remainder of this section specializes to the properties of the Legendre polynomials.

Theorem 10.9. *Schläfli's Integral: The integral representation given by eqs. (10.4.7) and (10.4.11) for the Hypergeometric function implies the following integral representation for the Legendre polynomials:*

$$P_\ell(z) = \frac{2^{-\ell}}{2\pi i}\int_\gamma dt\, \frac{(t^2-1)^\ell}{(t-z)^{\ell+1}} \tag{10.6.15}$$

in which the contour γ is a counterclockwise circle around the pole of the integrand at $t = z$.

Proof. The representation (10.4.11) is:

$$F(a,b;c;u) = C\int_\gamma ds\; s^{a-c}(s-1)^{c-b-1}(s-u)^{-a}. \tag{10.4.7}$$

Since $F(a,b;c;u) = F(b,a;c;u)$ we may switch $a \leftrightarrow b$ in this equation giving the following representation for $P_\ell(z)$:

$$P_\ell(z) = F\left[-\ell, 1+\ell; 1; \frac{1}{2}(1-z)\right] = C\int_\gamma ds\; s^\ell(s-1)^\ell\left(s-\frac{1}{2}+\frac{1}{2}z\right)^{-\ell-1}. \tag{10.6.16}$$

Choose the contour as indicated in the theorem. After the change of variables $t = 1-2s$ (10.6.16) becomes:

$$P_\ell(z) = C\int_\gamma dt\, \frac{(t^2-1)^\ell}{(t-z)^{\ell+1}}. \tag{10.6.17}$$

This contour satisfies the two conditions required of an integral representation quoted following eq. (10.4.11). All that remains is the evaluation of C. To do this evaluate both sides when $z = 1$, for which we know $P_\ell(1) = 1$, and perform the integration by residues:

$$1 = C \int_\gamma dt \, \frac{(t+1)^\ell}{t-1} = 2\pi i 2^\ell C. \tag{10.6.18}$$

This completes the proof. □

Another useful result for the purposes of evaluating integrals over combinations of Legendre polynomials is:

Theorem 10.10. *Rodriguez' Formula:*

$$P_\ell(z) = \frac{1}{2^\ell \ell!} \frac{d^\ell}{dz^\ell} \left[(z^2 - 1)^\ell \right]. \tag{10.6.19}$$

Proof. In order to prove this result, recall the Cauchy integral formula from the theory of complex variables:

$$f(z) = \frac{1}{2\pi i} \int_\gamma dt \, \frac{f(t)}{t-z} \tag{10.6.20}$$

in which γ is the contour encircling the point $t = z$ counterclockwise inside of which $f(t)$ is assumed analytic. Eq. (10.6.20) is proven by evaluating the integral by residues. Differentiating n times gives:

$$\frac{d^n f}{dz^n} = \frac{n!}{2\pi i} \int_\gamma dt \, \frac{f(t)}{(t-z)^{n+1}}. \tag{10.6.21}$$

The theorem then follows from a comparison of (10.6.21) with Schläfli's integral (10.6.15). □

The final useful property of the Legendre polynomials that is worth explicitly deriving is the generating function. The generating function is defined by the relation:

$$g(z,t) = \sum_{\ell=0}^{\infty} P_\ell(z) t^\ell. \tag{10.6.22}$$

The virtue of this function is its utility in proving properties for the all of the Legendre polynomials *at once*. Any property that can be proven for $g(z,t)$ for all t follows immediately for each term in its Taylor series in powers of t and so for each of the $P_\ell(z)$'s.

Theorem 10.11. *The generating function for the Legendre polynomials is given explicitly by:*

$$g(z,t) = (1 - 2zt + t^2)^{-1/2} = \sum_{\ell=0}^{\infty} P_\ell(z) t^\ell. \tag{10.6.23}$$

Proof. This expression may be derived by expressing the Legendre polynomials *via* Schläfli's integral and then directly evaluating the resulting sum and integral. Inserting (10.6.15) into the definition (10.6.22) gives:

$$g(z,t) = \frac{1}{2\pi i}\int_\gamma ds\, \frac{1}{s-z}\sum_{\ell=0}^{\infty}\left[\frac{(s^2-1)t}{2(s-z)}\right]^\ell . \tag{10.6.24}$$

The series is geometric and so may be easily evaluated: $\sum_0^\infty r^n = (1-r)^{-1}$ for $|r|<1$. The resulting formula for $g(z,t)$ is:

$$g(z,t) = -\frac{1}{i\pi t}\int_\gamma ds\, \frac{1}{s^2-1-2(s-z)/t} \tag{10.6.25}$$

for $|(s^2-1)t| < |2(s-z)|$.

The integral has the form $\int ds/[(s-s_+)(s-s_-)]$ where $s_\pm$ denote the roots of $s^2-1-2(s-z)/t$:

$$s_\pm = \frac{1}{t}[1 \pm \sqrt{1-2zt+t^2}]. \tag{10.6.26}$$

For small t these roots are located at $s_+ = 2/t - z + O(t)$ and $s_- = z + O(t)$. The pole at s_+ may be safely chosen outside of the integration contour while that at s_- must lie within. Evaluating the integral by residues at $s = s_-$ gives the desired result. □

10.7 Bessel Functions

Just as the Legendre functions are the archetype of the Hypergeometric equation, Bessel functions are the standard example of the Confluent Hypergeometric equation. We explore some of the properties of these functions in the present section.

Recall that Bessel's equation is:

$$y''(z) + \frac{1}{z}y'(z) + \left(1 - \frac{\nu^2}{z^2}\right)y(z) = 0. \tag{10.7.1}$$

The solutions are related to the Confluent Hypergeometric functions by eqs. (10.3.8) and(10.3.9). The conventional choice for the normalization of the solutions is:

$$y_1(z) = J_\nu(z) = \frac{1}{\Gamma(\nu+1)}\left(\frac{z}{2}\right)^\nu e^{-iz} M\left[\frac{1}{2}+\nu, 1+2\nu; 2iz\right] \tag{10.7.2}$$

for $\nu \neq -\frac{1}{2}, -1, -\frac{3}{2}, \ldots.$ The second, linearly independent solution is $y_2(z) = J_{-\nu}(z)$ provided that $\nu \neq \frac{1}{2}, 1, \frac{3}{2}, \ldots.$

The values taken by ν in the cases of physical interest vary. In the examples presented in chapter 5 Bessel's equation arose in two different contexts. In cylindrical coordinates the radial equation, (5.2.9), is Bessel's equation with $\nu = n$ an integer. In spherical coordinates the radial equation, (5.3.11), is Bessel's equation with $\nu = n + \frac{1}{2}$ a half-odd integer. The character of the two solutions is quite different because of this difference in the argument ν. These two cases are considered separately in the following paragraphs.

SPHERICAL BESSEL FUNCTIONS:

Consider first $\nu = n+\frac{1}{2}$: The solution to the radial equation (5.3.11) is $R(r) = y(\alpha r)/\sqrt{\alpha r}$ where α is a constant and $y(z)$ satisfies (10.7.1) with $\nu = n+\frac{1}{2}$. The first solution has a series form given by eq. (10.7.2) in terms of the Confluent Hypergeometric series with $a = n+1$ and $c = 2a = 2n+1$:

$$y_1(z) = J_{n+\frac{1}{2}}(z) = \frac{1}{\Gamma(n+\frac{3}{2})}\left(\frac{z}{2}\right)^{n+\frac{1}{2}} e^{-iz} M\left[n+1, 2n+2; 2iz\right] \tag{10.7.3}$$

for $n \neq -1, -2, ...$

The notation for these radial solutions is: $R(r) = j_n(\alpha r)$ with the spherical Bessel functions, $j_n(z)$, defined for $n \geq 0$ by:

$$j_n(z) \equiv \sqrt{\frac{\pi}{2z}} J_{n+\frac{1}{2}}(z). \tag{10.7.4}$$

The other solutions, that are not bounded at $z = 0$, are ($n \geq 0$):

$$n_n(z) = (-)^{n+1}\sqrt{\frac{\pi}{2z}} J_{-n-\frac{1}{2}}(z). \tag{10.7.5}$$

At first sight it would appear that, since c here takes on the integer values $2n+2$, the second solution to (10.7.3) is *not* given as in eq. (10.7.5) simply by reversing the sign of ν: *i.e.* $n \to -n-1$. This conclusion turns out in this case to be incorrect, however, because even though c takes integer values, it does so with $c = 2a$ with a also an integer. This fact is argued here in the form of the following:

Theorem 10.12. *The second solution to (10.7.1) with $\nu = n+\frac{1}{2}$ is given by:*

$$y_1(z) = J_{-n-\frac{1}{2}}(z) = \frac{1}{\Gamma(\frac{1}{2}-n)}\left(\frac{z}{2}\right)^{-n-\frac{1}{2}} e^{-iz} M\left[-n, -2n; 2iz\right]. \tag{10.7.6}$$

and the right-hand-side is well behaved even for positive integers n. Similarly, eq. (10.7.3) is really perfectly well-behaved for negative n.

Proof. To see why the series is not singular even though for $n \geq 0$ $c = -2n$ superficially appears to make some terms in the series (10.2.8) infinite consider the function $M(a, c; z)$ with $c = 2a$:

$$M(a, 2a; z) = 1 + \frac{1}{2}z + ... + \frac{1}{2}\frac{(a+1)...(a+k-1)}{(2a+1)...(2a+k-1)}\frac{z^k}{k!} + ... \tag{10.7.7}$$

Taking the limit that $a \to -n$ then gives for the $(k+1)$'th term:

$$\frac{(1-n)...(k-1-n)}{(1-2n)...(k-1-2n)}\frac{z^k}{k!} \tag{10.7.8}$$

The potential danger arises because the denominator can vanish when $k_{den} = 1 + 2n$. Notice however that the numerator may itself vanish when $k_{num} = 1 + n$. Now, since $n \geq 0$, k_{den} is larger or equal to k_{num} and they are equal only if $n = 0$. For $n > 0$ $k_{den} > k_{num}$ which implies that the series terminates, and since it does so before the terms with vanishing denominator have been reached each term in the sum is perfectly well-defined. *i.e.* for $n > 0$:

$$M(-n, -2n; z) = 1 + \frac{1}{2} \sum_{k=1}^{n+1} \frac{(1-n)...(k-1-n)}{(1-2n)...(k-1-2n)} \frac{z^k}{k!}. \tag{10.7.9}$$

In the case that $n = 0$ then $k_{den} = k_{num}$ so the zero of the denominator is cancelled by that of the numerator. This series is also well-defined:

$$M(0, 0; z) = 1 + \frac{1}{2} \sum_{k=1}^{\infty} \frac{z^k}{k!} = \frac{1}{2}(1 + e^z). \tag{10.7.10}$$

This establishes the theorem. $\square$

A corollary of this result is that the Bessel functions of half-integral argument, and hence the spherical Bessel functions, are expressable in terms of elementary functions: $J_{-\frac{1}{2}}(z) = \sqrt{\frac{2}{\pi z}} \cos(z)$ or, equivalently, $n_0(z) = \cos(z)/z$. Similarly $M(1, 2; z) = (e^z - 1)/z$ implies that $J_{\frac{1}{2}}(z) = \sin(z)/\sqrt{2\pi z}$ and $j_0(z) = \sin(z)/2z$. The functions with $n \neq 0$ are then related to these by differentiation as may be proven using recursion relations such as those discussed in section 10.5.

BESSEL FUNCTIONS OF INTEGER ARGUMENT:

The other case of physical interest is that in which $\nu = n$ is an integer. Then the Confluent Hypergeometric parameters for the two solutions $y_1(z)$ and $y_2(z)$ are given by $c = 2a = 1 \pm 2n$. Since c is an integer there is a potential difficulty in the definition of the first solution, (10.7.2), for n negative and in the second solution for n positive. Related to this is the possibility that the second solution is logarithmically singular at $z = 0$ in which case it does not admit such a series form.

Inspection of the Confluent Hypergeometric series shows that, unlike the situation just encountered for spherical Bessel functions, all of these potential difficulties really occur. This is because the series cannot terminate for the given values for a and so one is skewered upon the divergent $(2|n| + 2)$'th term of the series.

The regular solution (*i.e.* the one lacking a logarithmic singularity) for positive n is given by eq. (10.7.2) with $\nu = n$ and defines the Bessel function of non-negative integer argument $J_n(z)$. The Bessel function $J_{-n}(z)$ is similarly defined by the same equation with n replaced by $-n = |n|$ and so is not independent of $J_n(z)$. It is conventional to define $J_{-n}(z)$ with the following normalization relative to $J_n(z)$:

$$J_{-n}(z) = (-)^n J_n(z). \tag{10.7.11}$$

The utility of such a definition is that $J_m(z)$ so defined is regular at $z = 0$ regardless of the sign of of the integer m. The second solution is discussed later in this section.

We now turn to the derivation of some of the properties of Bessel functions.

Theorem 10.13. *Integral Representation: The Confluent Hypergeometric function has an integral representation, (10.4.12), that may be applied to the specific case of Bessel functions.*

$$J_\nu(z) = \frac{1}{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})}\left(\frac{z}{2}\right)^\nu \int_{-1}^{1} ds\ e^{izs}(1-s^2)^{\nu-\frac{1}{2}}. \tag{10.7.12}$$

Proof. The proof is a straightforward combination of eqs. (10.4.12) and (10.7.2). Recall that

$$M(a,c;z) = \frac{\Gamma(c)}{\Gamma(c-a)\Gamma(a)}\int_0^1 dt\ e^{zt}t^{a-1}(1-t)^{c-a-1} \tag{10.4.12}$$

for Re $c >$ Re $a > 0$. Using $c = 2a = \nu + \frac{1}{2}$ gives the following result:

$$J_\nu(z) = \frac{\Gamma(1+2\nu)}{\Gamma(\nu+1)[\Gamma(\nu+\frac{1}{2})]^2}\left(\frac{z}{2}\right)^\nu \int_0^1 dt\ e^{2izt}[t(1-t)]^{\nu-\frac{1}{2}}. \tag{10.7.13}$$

Next, the doubling formula, (6.5.19), implies that the combination $\Gamma(1+2\nu)/\Gamma(\frac{1}{2}+\nu)$ can be rewritten as:

$$\frac{\Gamma(1+2\nu)}{\Gamma(\frac{1}{2}+\nu)} = \frac{2^{2\nu}}{\sqrt{\pi}}\Gamma(1+\nu). \tag{10.7.14}$$

Finally the result (10.7.12) is obtained by changing variables using $s = 2t - 1$. □

This integral form allows the derivation of a simpler form for the Bessel series:

Theorem 10.14. *Bessel's series:*

$$J_\nu(z) = \sum_{n=0}^{\infty}(-)^n \frac{1}{\Gamma(\nu+n+1)n!}\left(\frac{z}{2}\right)^{2n+\nu}. \tag{10.7.15}$$

Proof. This series follows most simply from the integral representation (10.7.12) in which the exponential is Taylor expanded:

$$J_\nu(z) = \frac{1}{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})}\left(\frac{z}{2}\right)^\nu \sum_{k=0}^{\infty}\frac{(iz)^k}{k!}\int_{-1}^{1} ds\ s^k(1-s^2)^{\nu-\frac{1}{2}}. \tag{10.7.16}$$

The integral may be performed explicitly in terms of Euler's beta-function after the change of variables $u = 1 - s^2$. The required integral is:

$$\int_{-1}^{1} ds\ s^k(1-s^2)^{\nu-\frac{1}{2}} = \begin{cases} 0, & \text{if } k = 2n+1; \\ \frac{\Gamma(n+\frac{1}{2})\Gamma(\nu+\frac{1}{2})}{\Gamma(n+\nu+1)}, & \text{if } k = 2n\,. \end{cases} \tag{10.7.17}$$

Finally using the duplication formula in its form (10.7.14) one more time for the ratio $\Gamma(n+\frac{1}{2})/(2n)!$ gives the desired result. □

Another form for the integral representation often proves useful. This is given for integral $\nu = n$ by:

$$J_n(z) = \frac{1}{2\pi i} \int_\gamma \exp\left[\frac{1}{2} z \left(t - \frac{1}{t}\right)\right] \frac{1}{t^{n+1}}\, dt \tag{10.7.18}$$

in which γ traverses a small circle enclosing the singularity at $t = 0$ in the counterclockwise direction.

Proof. There are at least three ways to prove this result. The first is to directly evaluate the integral by residues after expanding the exponential in terms of its Taylor series. The resulting series expression for $J_n(z)$ agrees with the result (10.7.15) above.

Alternatively, operate on the right-hand-side of eq. (10.7.18) with the Bessel differential operator (10.7.1) and see that the result is the integral of a total derivative with respect to t that vanishes because the integration contour is closed. Since the integral converges and is analytic at $z = 0$ it must be proportional to $J_n(z)$ with a proportionality constant that can be determined by requiring that the n'th derivative at $z = 0$ agree with its known value as inferred from the first term of Bessel's series (10.7.15).

Method three starts with the integral representation (10.7.12) and the observation that the factor z^n can be rewritten as $(-id/ds)^n e^{izs}$ inside the integral. Then integrate the derivatives by parts off of the exponential and onto the other factor in the integrand. Finally change variables so $is = \frac{1}{2}(t - t^{-1})$ to get the result. $\square$

COROLLARY:

An immediate corollary is the extension of the integral representation (10.7.18) to non-integer ν (for Re $z > 0$):

$$J_\nu(z) = \frac{1}{2\pi i} \int_\gamma \exp\left[\frac{1}{2} z \left(t - \frac{1}{t}\right)\right] \frac{1}{t^{\nu+1}}\, dt \tag{10.7.19}$$

in which the integrand now has a branch point at $t = 0$ with branch cut chosen to lie along the negative real axis. The contour γ now starts at $t = -\infty - i\epsilon$ just below the cut and circles around the branch point at $t = 0$ to end up at $t = -\infty + i\epsilon$ just above the cut.

Proof. The proof follows approach 2 of the previous theorem. Direct differentiation shows that applying Bessel's equation to the right-hand-side gives the integral of a total derivative with respect to t. The resulting surface term vanishes since it involves a factor of $\exp[\frac{1}{2}z(t - t^{-1})]$ which vanishes as $t \to -\infty$ provided that Re $z > 0$. $\square$

This last form, (10.7.19), for $J_\nu(z)$ may be evaluated within the saddle-point approximation and so used to determine the asymptotic form for $J_\nu(z)$ at large z. There are two relevant saddle points—at $z = i$ and $z = -i$—that contribute to the result:

Theorem 10.15. *Asymptotic Form:*

$$J_\nu(z) = \frac{2}{\sqrt{\pi z}} \cos\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \left[1 + O(z^{-1})\right]; \qquad Re\, z > 0. \tag{10.7.20}$$

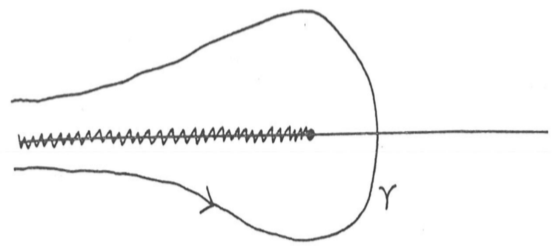

Figure 10.1. The Contour γ

This asymptotic behaviour motivates some conventional choices for the second solution to Bessel's equation that is linearly independent to $J_\nu(z)$. Consider therefore the following definitions:

Definition 10.2. *Hankel Functions:*

$$H_\nu^{(1)}(z) \equiv \frac{1}{\pi i} \int_{\gamma_1} \exp\left[\frac{1}{2} z \left(t - \frac{1}{t}\right)\right] \frac{1}{t^{\nu+1}}\, dt \tag{10.7.21}$$

and:

$$H_\nu^{(2)}(z) \equiv \frac{1}{\pi i} \int_{\gamma_2} \exp\left[\frac{1}{2} z \left(t - \frac{1}{t}\right)\right] \frac{1}{t^{\nu+1}}\, dt \tag{10.7.22}$$

define Hankel's functions of the first and second kinds respectively.

The contours in these definitions differ from that used in eq. (10.7.19). γ_1 consists of the contour starting at $t = 0$, moving out in the direction of positive Re t and then circling around to end up at $t = -\infty + i\epsilon$, just above the real axis. γ_2 is the reflection of γ_1 in the real axis traversed from $t = -\infty - i\epsilon$ to $t = 0$.

$H_\nu^{(1)}(z)$ and $H_\nu^{(2)}(z)$ are linearly independent solutions to Bessel's equation. They are both also linearly independent of $J_\nu(z)$. (The complete set of three solutions is, of course, linearly dependent.)

Proof. The proof is as for eq. (10.7.19). Substituting eqs. (10.7.21) and (10.7.22) into Bessel's equation (10.7.1) gives the integral of a total derivative that vanishes by virtue of the factor $\exp[\frac{1}{2} z(t - t^{-1})]$ in the integrand.

Linear independence is most easily established from the asymptotic forms for the Hankel functions, found from eqs. (10.7.21) and (10.7.22) by saddle-point integration:

$$H_\nu^{(1)}(z) = \frac{2}{\sqrt{\pi z}} \exp\left(iz - \frac{i\nu\pi}{2} - \frac{i\pi}{4}\right)\left[1 + O(z^{-1})\right] \tag{10.7.23}$$

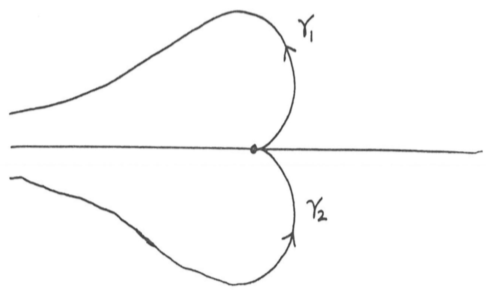

Figure 10.2. The Contours γ_1 and γ_2

and

$$H_\nu^{(2)}(z) = \frac{2}{\sqrt{\pi z}} \exp\left(-iz + \frac{i\nu\pi}{2} + \frac{i\pi}{4}\right) \left[1 + O(z^{-1})\right]. \tag{10.7.24}$$

Eqs. (10.7.23) and (10.7.24) establish the independence of the two Hankel functions since these equations are inconsistent with their being proportional to one another. Finally, since the contour γ in (10.7.19) is obtained when γ_2 and γ_1 are traversed in succession, it follows that Bessel's function, $J_\nu(z)$, is related to $H_\nu^{(1)}(z)$ and $H_\nu^{(2)}(z)$ by:

$$J_\nu(z) = \frac{1}{2}\left[H_\nu^{(1)}(z) + H_\nu^{(2)}(z)\right]. \tag{10.7.25}$$

□

The asymptotic forms (10.7.20), (10.7.23) and (10.7.24) and relation (10.7.25) show that Bessel functions and Hankel functions are related in the same way as the trigonometric function $\cos(z)$ is related to the exponentials $\exp(\pm iz)$. This suggests the definition of an alternative solution that is independent of $J_\nu(z)$ and which is the analogue of $\sin(z)$.

Definition 10.3. *Neumann function:*

$$N_\nu(z) \equiv -\frac{i}{2}\left[H_\nu^{(1)}(z) - H_\nu^{(2)}(z)\right]. \tag{10.7.26}$$

Inspection of eqs. (10.7.23) and (10.7.24) clearly shows that the asymptotic form for $N_\nu(z)$ is given by:

$$N_\nu(z) = \frac{2}{\sqrt{\pi z}} \sin\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \left[1 + O(z^{-1})\right]. \tag{10.7.27}$$

The two functions $J_\nu(z)$ and $N_\nu(z)$ form a basis of solutions to Bessel's O.D.E.. The same is true of the two Hankel functions. Which solutions are the more convenient to work with depends on the boundary conditions that are to be imposed. Bessel's and Neumann functions are naturally suited to boundary problems in which the solutions are required to be well-behaved at the origin since this immediately implies that the desired solution must be proportional to $J_\nu(z)$. Alternatively, if the solutions are known to behave like sine or cosine functions as $z \to \infty$ they are most naturally expressed in terms of either $J_\nu(z)$ or $N_\nu(z)$. If, on the other hand, solutions are required which behave like $\exp(iz)$ for large z then the basis of Hankel functions proves more convenient.

A further set of definitions is often convenient. Recall that Bessel's modified O.D.E.,

$$y''(z) + \frac{1}{z} y'(z) - \left(1 + \frac{\nu^2}{z^2}\right) y(z) = 0. \tag{10.7.28}$$

becomes Bessel's equation (10.7.1) after the change of variable $z \to iz$. It is conventional to denote the independent solutions to (10.7.28) as follows:

Definition 10.4. *Modified Bessel Functions:*

$$I_\nu(z) \equiv i^{-\nu} J_\nu(iz) \tag{10.7.29}$$

and

$$K_\nu(z) \equiv \frac{\pi}{2} i^{\nu+1} H_\nu^{(1)}(iz). \tag{10.7.30}$$

The powers of i appearing in these two definitions are chosen to ensure that the result is real when evaluated at real values of z. Whereas the convention (10.7.29) is fairly standard, the choice (10.7.30) is not. It is necessary to check carefully the definitions used in any particular reference.

This concludes our discussion of the properties of special functions.

11 Sturm-Liouville Problems

11.1 Introduction

In order to motivate the purpose of this chapter it is necessary to recap the story so far. The goal is and has always been the construction of solutions to boundary-value problems for second-order linear P.D.E.'s. This problem naturally comes in two parts. As was recognized in chapter 4, in a well-posed boundary-value problem the boundary information is not linear and homogeneous. It is usually true, however, that 'most' of the boundary information (*e.g.* that for all of the coordinates but one) is linear and homogeneous and it is only one piece of information (such as an intial condition) that is inhomogeneous and so uniquely specifies the solution.

The first step in solving a given boundary-value problem therefore consists of generating the general solutions to the linear and homogeneous part of the boundary-value problem. The key observation here is that the space of functions satisfying this sort of conditions forms a vector space and so it suffices to be able to construct a basis for this space.

Once such a basis is known the second step consists of the determination of the unique element of this vector space that satisfies the inhomogeneous part of the boundary information. The theory of the previous chapters consists of a set of techniques for constructing 'separated' solutions to the given O.D.E.'s. The problem of solving for separated solutions reduces to solving a set of linear second-order O.D.E.'s, which is a problem that is solved in some generality in chapters 8, 9 and 10. The big drawback of this technique is that almost none of the solutions to the homogeneous boundary-value problem can be expected to really have this separated form.

The crucial missing piece in the argument is therefore the demonstration that although the general solution is not separated, it is often possible to construct a *basis* of separated solutions. This implies that the general solution can be constructed as a linear combination of separated solutions and hence that the separation technique can really be used to generate general solutions to many problems. It is this missing step that is intended to be filled in the present chapter. A secondary outcome of this chapter is the development of explicit expressions that determine the unique linear combination of basis solutions that satisfies a given inhomogeneous boundary condition.

11.2 Some Linear Algebra

The entire presentation is aimed at demonstrating a parallel between the types of differential equations being solved and the familiar eigenvalue problem from linear algebra. This section is meant to be a short reminder of the relevant properties of finite-dimensional eigenvalue problems.

The O.D.E.'s solved in the previous chapters had the following generic form:

$$\mathcal{A}y(z) = \lambda y(z) \tag{11.2.1}$$

in which $\mathcal{A}$ is some second-order differential operator and λ is an undetermined separation constant that arose in the process of separating the original P.D.E. into a set of independent O.D.E.'s. It was generally found that the (homogeneous) boundary conditions imposed determined that (11.2.1) had no solutions unless λ took on specific values λ_n.

A simple example of this pattern is given by the function $X(\xi)$ in cylindrical coordinates or $\Phi(\phi)$ in spherical coordinates. In both cases the boundary-value problem of interest was of the form (11.2.1) with $\mathcal{A} = d^2/dz^2$ and homogeneous boundary condition $y(z+2\pi) = y(z)$. Solutions only existed for $\lambda_n = -n^2$, with n an integer, and the corresponding solutions were $y_n(z) = e^{inz}$.

All of this smacks of an eigenvalue problem.

Consider, then, an n-dimensional vector space, V, and an n-by-n hermitian (or self-adjoint) matrix $\mathbf{A} = \mathbf{A}^\dagger$.

Definition 11.1. *The eigenvalue problem for the N-by-N matrix $\mathbf{A}$ is given by:*

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{11.2.2}$$

in which both the nonzero eigenvectors, $\mathbf{v}$, and the constant eigenvalues, λ, are to be determined for given $\mathbf{A}$.

Writing eq. (11.2.2) as $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = 0$ shows that it has solutions only for eigenvalues that satisfy $P(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Since $P(\lambda)$ is an N-th degree polynomial in λ it has N complex roots, λ_k, with $k = 1, ..., N$. The corresponding eigenvectors are denoted $\mathbf{v}_k$. They satisfy the following general properties whose analogues for differential equations are crucial to our argument:

Theorem 11.1. *If $\mathbf{A}$ is hermitian then its eigenvalues, λ_k, are real.*

Proof. To prove this consider the multiplying the eigenvalue equation (11.2.2) on the left by $\mathbf{v}^\dagger$. The result is:

$$\mathbf{v}^\dagger\mathbf{A}\mathbf{v} = \lambda\mathbf{v}^\dagger\mathbf{v} \tag{11.2.3}$$

The observation that $(\mathbf{v}^\dagger\mathbf{v})^* = (\mathbf{v}^\dagger\mathbf{v})$ is real, as is

$$(\mathbf{v}^\dagger\mathbf{A}\mathbf{v})^* = (\mathbf{v}^\dagger\mathbf{A}^\dagger\mathbf{v}) = (\mathbf{v}^\dagger\mathbf{A}\mathbf{v}) \tag{11.2.4}$$

implies the desired reality of the eigenvectors:

$$\lambda = \lambda^*. \tag{11.2.5}$$

□

Theorem 11.2. *Eigenvectors, $\mathbf{v}_k$ and $\mathbf{v}_l$, for differing eigenvalues, $\lambda_k \neq \lambda_l$, are orthogonal:* $\mathbf{v}_k^\dagger\mathbf{v}_l = 0$.

Proof. Consider the eigenvalue equation, (11.2.1), for the two given eigenvectors:

$$\mathbf{A}\mathbf{v}_k = \lambda_k \mathbf{v}_k \tag{11.2.6}$$

and

$$\mathbf{A}\mathbf{v}_l = \lambda_n \mathbf{v}_l. \tag{11.2.7}$$

Multiply eq. (11.2.6) on the left by $\mathbf{v}_l^\dagger$ and eq. (11.2.7) by $\mathbf{v}_k^\dagger$. Taking the difference between one of these equations and the complex conjugate of the other gives the results:

$$\begin{aligned} 0 &= (\mathbf{v}_l^\dagger [\mathbf{A}^\dagger - \mathbf{A}]\mathbf{v}_k) \\ &= [(\mathbf{A}\mathbf{v}_l)^\dagger \mathbf{v}_k] - (\mathbf{v}_l^\dagger \mathbf{A}\mathbf{v}_k) \\ &= (\lambda_l - \lambda_k)(\mathbf{v}_l^\dagger \mathbf{v}_k) \end{aligned} \tag{11.2.8}$$

which, since $\lambda_l \neq \lambda_k$ by assumption, implies the result. □

Theorem 11.3. *The eigenvectors, $\mathbf{v}_k$, of a hermitian matrix, $\mathbf{A}$, are complete. That is, they form a basis for the vector space, V, which ensures that any vector can be written as: $\mathbf{w} = \sum_k c_k \mathbf{v}_k$.*

Proof. If all of the eigenvalues of $\mathbf{A}$ are distinct then this conclusion is straightforward. This is because there are N distinct eigenvalues, each of which corresponds to a distinct eigenvector. Since the eigenvalues are distinct, all of these eigenvectors are mutually orthogonal: $(\mathbf{v}_k^\dagger \mathbf{v}_l) = 0$ for all $l \neq k$. Being orthogonal they must be linearly independent. (This last result follows by assuming that $C_1\mathbf{v}_1 + ...C_N\mathbf{v}_N = 0$ and then multiplying on the left by $\mathbf{v}_k^\dagger$ for each k in succession. Since $(\mathbf{v}_k^\dagger \mathbf{v}_k) \neq 0$ it follows that each of the C_k's vanishes.) Now, as proven in section 3.2, *any* set of N linearly-independent vectors in an N-dimensional vector space forms a basis and so the argument is complete.

The only remaining point is what happens if not all of the eigenvalues are distinct, *i.e.* the polynomial $P(\lambda)$ defined following eq. (11.2.2) has degenerate roots. The set of eigenvectors $\mathbf{v}_k$ that satisfy eq. (11.2.1) for the same eigenvalue λ forms a subspace of the full vector space V called the *eigenspace* corresponding to the eigenvalue λ. The question to be understood is what the dimension of this subspace is. It is a standard theorem of linear algebra, quoted here without proof, that the eigenspace of an n-fold degenerate eigenvalue, λ, has dimension n and so always contains n linearly-independent eigenvectors. This guarantees that even when some eigenvalues are degenerate there are N independent eigenvectors that must span V. □

Given a basis of eigenvectors it is desirable to be able to explicitly compute the linear combination that corresponds to a given vector. For these purposes it is convenient to work with a basis that is *orthonormal*, *i.e.* orthogonal and normalized to unity: $\mathbf{v}_k^\dagger \mathbf{v}_l = \delta_{kl}$. ($\delta_{kl}$ is the Kronecker delta.) In this case the claim now demonstrated is that it is always possible to explicitly construct a basis of eigenvectors that are mutually orthogonal. The method employed is known as *Gram-Schmidt* orthogonalization.

Theorem 11.4. *Gram-Schmidt Orthogonalization: Suppose that* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *are any two linearly independent vectors. Then the new combination:*

$$\begin{aligned}\mathbf{w}_1 &= \mathbf{v}_1 \\ \mathbf{w}_2 &= \mathbf{v}_2 - \frac{(\mathbf{v}_1^\dagger \mathbf{v}_2)}{(\mathbf{v}_1^\dagger \mathbf{v}_1)}\mathbf{v}_1\end{aligned} \tag{11.2.9}$$

is orthogonal: $\mathbf{w}_1^\dagger \mathbf{w}_2 = 0$.

Proof. The proof consists of explicitly multiplying out the product: $\mathbf{w}_1^\dagger \mathbf{w}_2 = 0$ using the definition (11.2.9) for the vectors $\mathbf{w}_k$. □

This process may be repeated as many times as is necessary to convert a given basis of functions, $\mathbf{v}_k$, into an orthogonal basis $\mathbf{w}_k$. Dividing each element $\mathbf{w}_k$ of the orthogonal basis by its length then gives an orthonormal basis $\mathbf{e}_k = \mathbf{w}_k / \sqrt{\mathbf{w}_k^\dagger \mathbf{w}_k}$.

Theorem 11.5. *Any vector,* $\mathbf{w}$*, in* V *can be written in terms of an orthonormal basis as:*

$$\mathbf{w} = \sum_{k=1}^{N} c_k \mathbf{e}_k \tag{11.2.10}$$

with the constants c_k *given explicitly by*

$$c_k = \mathbf{e}_k^\dagger \mathbf{w}. \tag{11.2.11}$$

Proof. The existence and uniqueness of the expansion follows from the fact that the $\mathbf{e}_k$'s are assumed to form a basis for the vector space. Formula (11.2.11) for the coefficients is obtained by substituting eq. (11.2.10) into $\mathbf{e}_k^\dagger \mathbf{w}$. □

11.3 Infinite-Dimensional Generalizations

We wish to be able to prove the analogue of the results listed above in the case where the vector space V is the space of functions satisfying a set of linear and homogeneous boundary conditions. The idea is that differential operators are the analogues of matrices acting in these spaces. The O.D.E.'s solved in previous chapters are the analogues of the eigenvalue problem (11.2.1) in which the unknown separation constants are the eigenvalues and the explicit solutions, $y(z)$, constructed are the *eigenfunctions*, *i.e.* the analogues of eigenvectors. We wish to prove that the analogue of the completeness of these eigenfunctions in this vector space as well as to be able to explicitly compute the analogue of eq. (11.2.11) that determines the coefficients appearing in the eigenfunction expansion of a given element of V.

To do so we abstract those properties of the 'dot product' $\mathbf{v}^\dagger \mathbf{w}$ and 'hermitian matrix' that are responsible for the proofs of the above results but which may be applied to our space of functions. Consider therefore the following definitions:

Definition 11.2. *Suppose f and g are arbitrary elements of some vector space V. Then any rule that generates a complex number, (f,g), from these elements and satisfies the following three properties is called an inner product.*

1.
$$(f, \alpha g + \beta h) = \alpha (f,g) + \beta (f,h) \tag{11.3.1}$$
for any vectors f, g and h in V and complex numbers α and β.

2.
$$(f,g)^* = (g,f) \tag{11.3.2}$$
for all f and g.

3.
$$\|f\|^2 \equiv (f,f) \geq 0 \qquad \textit{and} \qquad (f,f) = 0 \Longrightarrow f = 0. \tag{11.3.3}$$

These properties are satisfied by the old faithful definition $(\mathbf{v}, \mathbf{w}) = \mathbf{v}^\dagger \mathbf{w}$ in the special case of a finite-dimensional vector space. The claim is that these are the essential properties for the purposes of establishing the analogues of the results of the previous section. A more interesting realization of this definition is furnished by the following:

EXAMPLE:

Consider the vector space, V, of (piecewise smooth, say) functions defined on some interval, $x \in [a,b]$, of the real line and that satisfy some linear and homogeneous boundary condition at the endpoints a and b. Then an inner product for this vector space is given by:

$$(f,g)_w \equiv \int_a^b f^*(x) g(x) w(x) \, dx \tag{11.3.4}$$

in which $w(x)$ is some given real, non-negative function that can only vanish at isolated points in the interval $[a,b]$. Notice that $w(x)$ is a function that is specified once and for all and that it need not itself be an element of the vector space under consideration. It is easy to check that definition (11.3.4) satisfies all of the requirements listed in eqs. (11.3.1), (11.3.2) and (11.3.3).

Definition 11.3. *A differential operator, $\mathcal{A}$, that takes V to itself is said to be self-adjoint with respect to an inner product, $(f,g)_w$ if it satisfies:*

$$(f, \mathcal{A} g)_w = (\mathcal{A} f, g)_w \tag{11.3.5}$$

for all f and g in V.

CAVEAT! It should be noted that for infinite-dimensional vector spaces there are distinctions of definition between self-adjoint and hermitian operators and that the definition given here ignores all of these niceties. The distinctions have to do with what the image of V is under the action of the operator. That is, it is often the case that $\mathcal{A} f$ may not satisfy the boundary or smoothness conditions defining V even though f itself does. All these subtleties are brazenly ignored in what follows.

Definition 11.4. *An eigenvalue problem for a self-adjoint operator, $\mathcal{A}$, is a differential equation:*

$$\mathcal{A}f = \lambda f \tag{11.3.6}$$

in which λ and f are to be determined given $\mathcal{A}$.

Just as was the case for finite-dimensional spaces, (11.3.6) in general admits solutions only for some λ_n's. The eigenvalues are denoted by λ_n, and the corresponding eigenfunctions are denoted f_n.

The utility of these definitions is in the fact that they imply the analogues of the results of the previous section for these differential eigenvalue problems. We repeat those results again here in terms of these new definitions.

Theorem 11.6. *If $\mathcal{A}$ is self-adjoint then its eigenvalues, λ, are real.*

Proof. As before consider $(f, \mathcal{A}f)^*$. The definitions of inner product and self-adjointness imply: $(f, \mathcal{A}f)^* = (\mathcal{A}f, f) = (f, \mathcal{A}f)$, so $(f, \mathcal{A}f)$ is real. Now, if f is an eigenfunction of $\mathcal{A}$ then $(f, \mathcal{A}f) = \lambda(f, f)$. Since (f, f) is real by definition, this last equation implies the reality of the eigenvalue λ. □

Theorem 11.7. *Eigenfunctions, f_k and f_l, for differing eigenvalues, $\lambda_k \neq \lambda_l$, are orthogonal: $(f_k, f_l) = 0$.*

Proof. Consider in this case:

$$\begin{aligned} 0 &= (\mathcal{A}f_k, f_l) - (f_k, \mathcal{A}f_l) \\ &= (\lambda_k - \lambda_l)(f_k, f_l) \end{aligned} \tag{11.3.7}$$

and use $\lambda_l \neq \lambda_k$ to get the desired result. □

Definition 11.5. *An orthonormal set of elements, f_k, of V satisfy the following definition: $(f_k, f_l) = \delta_{kl}$.*

Theorem 11.8. *Gram-Schmidt Orthogonalization: Given any basis, f_k, of V an orthonormal basis, v_k, may be constructed by the Gram-Schmidt construction.*

Proof. The construction is identical to the finite-dimensional case. Given two linearly independent elements, f_1 and f_2, of V an orthonormal pair can be constructed *via* the following two steps. First define the orthogonal combinations:

$$u_1 = f_1 \tag{11.3.8}$$

and

$$u_2 = f_2 - \frac{(f_1, f_2)}{(f_1, f_1)} f_1. \tag{11.3.9}$$

These may then be normalized by the choice $v_k = u_k/\sqrt{(u_k, u_k)}$. □

Theorem 11.9. *Completeness (A Cartoon): The eigenfunctions, v_k, of a self-adjoint operator span the space, V, of piecewise smooth functions in which the eigenvalue problem was posed. More precisely any piecewise smooth function, f, in V may be represented uniquely by the infinite sum:*

$$f(x) = \sum_{k=1}^{\infty} c_k v_k(x) \tag{11.3.10}$$

In general the sum on the right-hand-side converges to the function on the left-hand-side in the sense of *convergence in the mean*:

$$\lim_{N\to\infty} \|f - \sum_{k=1}^{N} c_k v_k\|^2 = 0. \tag{11.3.11}$$

Notice that this is a weaker notion of convergence than pointwise convergence so that it does not follow that the sum on the right and the function on the left are equal for all x. This failure of the sum to converge to the value of the function is generally limited to the points of discontinuity of $f(x)$, and is often referred to as *Gibb's phenomenon*. For neighbourhoods in which $f(x)$ and $v_k(x)$ are smooth, on the other hand, this theorem can be strengthened to the statement that the right-hand-side converges uniformly to the left-hand-side. No attempt will be made to prove this theorem.

Theorem 11.10. *If $v_k(x)$ are chosen to be a set of orthonormal eigenfunctions of a self-adjoint operator, $\mathcal{A}$, in terms of which a general element of V may be expressed by eq. (11.3.10), then the coefficients, c_k, of this expansion are given by:*

$$c_k = (v_k, f)_w = \int_a^b v_k^*(x) f(x) w(x)\, dx \tag{11.3.12}$$

in which $(v_k, f)_w$ is the inner product with respect to which $\mathcal{A}$ is self-adjoint.

These properties are illustrated in several examples at the end of the next section.

11.4 Self-Adjoint Operators

Before turning to examples of the application of these results to the differential equations and problems of interest, one more issue must be addressed. All of these theorems apply to the eigenfunctions of self-adjoint operators so a method is necessary to systematically put a given differential equation into self-adjoint form. Once this has been done all of the results of the previous section may be taken over as a body.

Such a method is given by the following observation: Suppose the linear, second-order O.D.E. of interest may be put into the following *Sturm-Liouville* form:

$$\mathcal{L}y(x) = \lambda w(x) y(x) \tag{11.4.1}$$

in which λ is an undetermined constant, $w(x)$ is a real, non-negative function that vanishes only at isolated points within the interval $[a, b]$ and where the differential operator, $\mathcal{L}$, is given by:

$$\mathcal{L}y(x) = \frac{d}{dx}\left[A(x)\frac{dy}{dx}\right] + B(x)y \tag{11.4.2}$$

The O.D.E. may be thought of as an eigenvalue problem for the differential operator $\mathcal{A} = \frac{1}{w(x)}\mathcal{L}$.

Theorem 11.11. *This differential operator is self-adjoint with respect to the inner product* $(f, g)_w$ *defined by the function* $w(x)$ *provided that the following two properties hold:*

1. $$A(x) = A^*(x) \qquad \text{and} \qquad B(x) = B^*(x). \tag{11.4.3}$$

2. $$\left[A(x)\left(f^*\frac{dg}{dx} - \frac{df^*}{dx}g\right)\right]_a^b = 0 \tag{11.4.4}$$

   for all f *and* g *in* V.

Proof. We require the conditions under which $(\mathcal{A}f, g)_w = (f, \mathcal{A}g)_w$. Using the definition of $\mathcal{A}$ given in the statement of the theorem and eq. (11.3.4) for the inner product gives:

$$(\mathcal{A}f, g)_w - (f, \mathcal{A}g)_w = \int_a^b \left[(\mathcal{L}f)^*g - f^*(\mathcal{L}g)\right]\, dx \tag{11.4.5}$$

which after integrating by parts becomes:

$$\begin{aligned}(\mathcal{A}f, g)_w - (f, \mathcal{A}g)_w = &\int_a^b \left[(A - A^*)\frac{df^*}{dx}\frac{dg}{dx} - (B - B^*)f^*g\right]\, dx + \\ &+ \left[A^*\frac{df^*}{dx}g - Af^*\frac{dg}{dx}\right]_a^b\end{aligned} \tag{11.4.6}$$

This clearly vanishes once conditions (11.4.3) and (11.4.4) are satisfied. □

It should be noticed at this point that the assumption that the O.D.E. could be put into Sturm-Liouville form, (11.4.1) and (11.4.2), is very mild. To see this consider an arbitrary linear, second-order O.D.E. of the form:

$$a(x)\frac{d^2y}{dx^2} + b(x)\frac{dy}{dx} + c(x)y - \lambda w(x)y = 0. \tag{11.4.7}$$

This is put into Sturm-Liouville form by the change of variables:

$$y(x) = a(x)\exp\left[-\int^x \frac{b(u)}{a(u)}\, du\right] v(x). \tag{11.4.8}$$

Consider the following

11.5 Examples

1. Harmonic Oscillator Equation.

$$y''(z) + \alpha y(z) = 0. \tag{11.5.1}$$

Separation of variables gives this O.D.E. in several instances. Two common forms for the boundary conditions are:

$$(i) \text{ Periodicity:} \qquad z \in [0, 2\pi] \qquad \text{with} \qquad y(z + 2\pi) = y(z) \tag{11.5.2}$$

or:

$$(ii) \text{ Dirichlet conditions:} \qquad z \in [0, L] \qquad \text{with} \qquad y(0) = y(L) = 0 \tag{11.5.3}$$

These boundary conditions are both linear and homogeneous and so define the space of functions $V_{(i)}$ and $V_{(ii)}$ respectively.

This O.D.E. is already in Sturm-Liouville form. That is comparison of eq. (11.5.1) with eq. (11.4.1) and (11.4.2) gives the identification:

$$\mathcal{A} = \mathcal{L} = \frac{d^2}{dz^2} \tag{11.5.4}$$

so $A(z) = w(z) = 1$, $B(z) = 0$ and $\lambda = -\alpha$. Clearly both $A(z)$ and $B(z)$ are real. The second criterion for self-adjointness, eq. (11.4.4), becomes:

$$\left[\frac{df^*}{dx} g - f^* \frac{dg}{dx} \right]_a^b = 0 \tag{11.5.5}$$

which is satisfied for both choices (i) and (ii) of boundary conditions. The operator (11.5.4) is therefore self-adjoint in both $V_{(i)}$ and $V_{(ii)}$ with respect to the corresponding inner product:

$$(f, g) = \int_a^b f^*(x) g(x)\, dx. \tag{11.5.6}$$

The solutions to the O.D.E. have been found in previous sections and give the eigenvalues and normalized eigenfunctions as:

$$\text{Case } (i) \qquad \lambda_n = -n^2; \qquad v_n(z) = \frac{1}{\sqrt{2\pi}} e^{inz}; \qquad n = 0, \pm 1, ... \tag{11.5.7}$$

or

$$\text{Case } (ii) \qquad \lambda_n = -\left(\frac{\pi n}{L}\right)^2; \qquad v_n(z) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi n z}{L}\right); \qquad n = 1, 2, ... \tag{11.5.8}$$

Completeness implies that for any function in $V_{(i)}$:

$$\text{Case } (i) \qquad f(z) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n e^{inz} \tag{11.5.9}$$

with

$$c_n = (v_n, f) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{-inz} f(z)\, dz. \qquad (11.5.10)$$

This is recognized as a Fourier series. For $f(z) \in V_{(ii)}$:

$$\text{Case } (ii) \qquad f(z) = \sqrt{\frac{2}{L}} \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi n z}{L}\right) \qquad (11.5.11)$$

with

$$c_n = (v_n, f) = \sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{\pi n z}{L}\right) f(z)\, dz. \qquad (11.5.12)$$

2. Associated Legendre Equation.

$$\frac{d}{dz}\left[(1-z^2)\frac{dy}{dz}\right] - \alpha y - \frac{m^2}{1-z^2} y = 0 \qquad (11.5.13)$$

This equation arises when separating variables in spherical polar coordinates with the following boundary conditions:

$$z \in [-1, 1] \qquad \text{with} \qquad \lim_{z \to \pm 1} |y(z)| < \infty. \qquad (11.5.14)$$

These are linear homogeneous boundary conditions and so define the space of functions V.

Eq. (11.5.13) is already in Sturm-Liouville form so a comparison of with eqs. (11.4.1) and (11.4.2) implies:

$$\mathcal{A} = \frac{d}{dz}\left[(1-z^2)\frac{d}{dz}\right] - \frac{m^2}{1-z^2} \qquad (11.5.15)$$

so $A(z) = 1 - z^2$, $B(z) = -m^2/(1-z^2)$, $w(z) = 1$ and $\lambda = \alpha$. $A(z)$ and $B(z)$ are real. The second criterion for self-adjointness is eq. (11.4.4):

$$\left[(1-z^2)\left(\frac{df^*}{dx} g - f^* \frac{dg}{dx}\right)\right]_{-1}^{1} = 0 \qquad (11.5.16)$$

which is satisfied in this case because the function $A(z)$ vanishes at the endpoints $z = \pm 1$ while f and g remain bounded there. The operator (11.5.15) is therefore self-adjoint in V with respect to the corresponding inner product:

$$(f, g) = \int_{-1}^{1} f^*(z) g(z)\, dz. \qquad (11.5.17)$$

The solutions to the O.D.E. have been found in previous sections and give the eigenvalues and normalized eigenfunctions as:

$$\lambda_n = -\ell(\ell+1); \qquad v_\ell(z) = C_\ell {P_\ell}^m(z). \qquad (11.5.18)$$

Orthogonality of the eigenfunctions reads:

$$\int_{-1}^{1} (P_\ell{}^m)^*(z) P_r{}^s(z)\, dz = 0 \qquad \text{for} \qquad \ell \neq r \tag{11.5.19}$$

and the normalization condition determines the constant C_ℓ to be:

$$C_\ell^{-2} = \int_{-1}^{1} |P_\ell{}^m(z)|^2\, dz = \left(\frac{2}{2\ell+1} \frac{(\ell+m)!}{(\ell-m)!} \right). \tag{11.5.20}$$

Completeness now expresses the statement that any piecewise smooth member of V has the unique representation:

$$f(z) = \sum_{\ell=|m|}^{\infty} c_\ell \sqrt{\frac{2\ell+1}{2} \frac{(\ell-m)!}{(\ell+m)!}}\, P_\ell{}^m(z) \tag{11.5.21}$$

with

$$c_\ell = \sqrt{\frac{2\ell+1}{2} \frac{(\ell-m)!}{(\ell+m)!}} \int_{-1}^{1} (P_\ell{}^m(z))^* f(z)\, dz. \tag{11.5.22}$$

3. Spherical Harmonics.

   The last two examples frequently appear in the following product form:

$$Y_{\ell m}(\theta, \phi) = \Theta_{\ell,m}(\theta) \Phi_m(\phi) = (-)^m \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}}\, P_\ell{}^m(\cos\theta) e^{im\phi}. \tag{11.5.23}$$

The previous two examples therefore imply that the $Y_{\ell m}$'s are simultaneous eigenfunctions of the following two differential operators:

$$\frac{1}{\sin\theta} \left[\frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial Y_{\ell m}}{\partial\theta} \right) + \frac{1}{\sin\theta} \frac{\partial^2 Y_{\ell m}}{\partial\phi^2} \right] = -\ell(\ell+1) Y_{\ell m} \tag{11.5.24}$$

and

$$-i \frac{\partial Y_{\ell m}}{\partial\phi} = m Y_{\ell m}. \tag{11.5.25}$$

In quantum mechanics these two conditions correspond to the statement that the $Y_{\ell m}$'s represent states of definite angular momentum. Orthogonality is the statement:

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta\ \sin\theta [Y_{\ell m}(\theta,\phi)]^* Y_{rs}(\theta,\phi) = \delta_{\ell r}\delta_{ms} \tag{11.5.26}$$

while completeness states that:

$$f(\theta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell m} Y_{\ell m}(\theta,\phi) \tag{11.5.27}$$

with

$$c_{\ell m} = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\ \sin\theta [Y_{\ell m}(\theta,\phi)]^* f(\theta,\phi). \tag{11.5.28}$$

4. Bessel's Equation

$$z^2 y''(z) + z y'(z) + (\alpha^2 z^2 - m^2) y(z) = 0. \tag{11.5.29}$$

This equation arises in cylindrical coordinates with the constant α undetermined. A typical boundary condition might be:

$$z \in [0, a] \qquad \text{with} \qquad y(0) < \infty \qquad \text{and} \qquad y(a) = 0 \tag{11.5.30}$$

which are linear and homogeneous boundary conditions and define the vector space of functions V.

The equation may be put into Sturm-Liouville form by dividing through by z after which a comparison of with eq. (11.4.1) and (11.4.2) implies:

$$\mathcal{A} = \frac{d}{dz}\left[z\frac{d}{dz}\right] - \frac{m^2}{z} \tag{11.5.31}$$

so $A(z) = z$, $B(z) = -m^2/z$, $w(z) = z$ and $\lambda = -\alpha^2$. $A(z)$ and $B(z)$ are real. The boundary term in the self-adjointness test is eq. (11.4.4):

$$\left[z\left(\frac{df^*}{dx}g - f^*\frac{dg}{dx}\right)\right]_0^a = 0 \tag{11.5.32}$$

which is satisfied in this case for different reasons at each endpoint. For $z = 0$ it is because the function $A(z)$ vanishes while f and g remain bounded. At $z = a$ the functions f and g all vanish and so ensure that this contribution dies.

The operator (11.5.31) is therefore self-adjoint in V with respect to the corresponding inner product:

$$(f, g) = \int_0^a f^*(z) g(z) z \, dz. \tag{11.5.33}$$

The solutions to the O.D.E. have been found in previous sections. The solutions to (11.5.29) are

$$y(z) = C_1 J_m(\alpha z) + C_2 N_m(\alpha z) \tag{11.5.34}$$

The condition that the solution be regular at the origin implies $C_2 = 0$ and the boundary condition at $z = a$ determines the eigenvalue:

$$J_m(\alpha a) = 0 \tag{11.5.35}$$

This eigenvalue equation has an infinite number of roots as may be seen from the asymptotic form for the Bessel function, eq. (10.7.19). Numbering these by $n = 1, 2, ...$ allows the eigenvalues to be written as $\lambda_n = -\alpha_{mn}^2$ where the index m in α_{mn} is a reminder of which Bessel function appeared in the condition (11.5.35). The corresponding normalized eigenfunctions are:

$$v_n(z) = C_n J_m(\alpha_{mn} z) \tag{11.5.36}$$

with the normalization constant given by:

$$C_n^{-2} = \int_0^a |J_m(\alpha_{mn} z)|^2 z \, dz \tag{11.5.37}$$

so completeness implies that the piecewise smooth elements of V have the expansion:

$$f(z) = \sum_{n=1}^{\infty} c_n C_n J_m(\alpha_{mn} z) \tag{11.5.38}$$

with

$$c_n = C_n \int_0^a [J_m(\alpha_{mn} z)]^* f(z) z \, dz \tag{11.5.39}$$

12 The Grand Synthesis

We are finally in a position to put together all of the pieces accumulated in the preceding chapters. The best way to do so is to briefly state the procedure and then to work through a series of examples.

The general procedure for solving boundary-value problems for partial differential equations by separation of variables is alluded to in the simple example of section 5.1 and is as follows:

1. Separate variables to convert the given partial differential equation into a series of ordinary differential equations containing a number of undetermined separation constants.

2. Solve these equations, for example by the method of power-series solutions.

3. Identify which of the boundary conditions are linear and homogeneous and are also of a form that may be imposed separately on the separated ordinary differential equations. Some of the resulting O.D.E. boundary-value problems are eigenvalue problems of the Sturm-Liouville form and others may not be.

4. Impose these boundary conditions to eliminate as many integration constants as possible and, for the Sturm-Liouville problems, to determine what eigenvalues the separation constants may take.

5. Use the completeness theorem for the eigenfunctions of the Sturm-Liouville problems to expand the *general* solution to the P.D.E. together with the homogeneous boundary-value information in terms of separated solutions.

6. Finally, use formula (11.3.12) to find the specific solution to the P.D.E. that satisfies the remaining inhomogeneous boundary information.

Next, let's apply this reasoning to solve the following illustrative boundary-value problems:

12.1 Example I: Laplace's Equation

For the first problem consider a boundary-value problem for an elliptic differential equation.

Problem 12.1. *Solve for the electrostatic potential, $\varphi(r, \theta, \phi)$, inside a spherical cavity of radius a whose walls are held at a fixed potential $V(\theta, \phi)$.*

The differential equation governing the electrostatic potential is:

$$\nabla^2 \varphi = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2 \psi}{\partial \phi^2} = 0. \tag{12.1.1}$$

and the boundary conditions are that $\varphi(r, \theta, \phi)$ is nonsingular throughout the interior of the sphere and $\varphi(a, \theta, \phi) = V(\theta, \phi)$.

The first step is to generate the separated solutions $\varphi(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi)$ by separating variables in spherical coordinates. The separated equations together with their boundary conditions are:

1. $\Phi(\phi)$:

   This factor satisfies the O.D.E. (5.3.6):

   $$\Phi''(\phi) = \alpha\Phi(\phi) \tag{12.1.2}$$

   together with the boundary condition that $0 \leq \phi \leq 2\pi$ with the homogeneous condition $\Phi(\phi + 2\pi) = \Phi(\phi)$.

   The solutions to this equation are:

   $$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \tag{12.1.3}$$

   with m an integer, implying that the separation constant is $\alpha = -m^2$. As was noted in the previous chapter, eq. (12.1.2) is a Sturm-Liouville-type equation with eigenvalue $\lambda_m = -m^2$. The corresponding inner product is:

   $$(f, g) = \int_0^{2\pi} f^*(\phi)g(\phi)\, d\phi \tag{12.1.4}$$

   and the eigenfunctions are orthonormal with respect to this inner product when normalized as in eq. (12.1.3).

2. $\Theta(\theta)$:

   The differential equation to be solved is the associated Legendre equation (5.3.10):

   $$\frac{d}{dx}\left[(1-x^2)\frac{d\Theta}{dx}\right] - \left(\beta + \frac{m^2}{1-x^2}\right)\Theta = 0 \tag{5.3.10}$$

   where $x = \cos\theta$. The boundary condition is that Θ should not diverge for any θ including the endpoints at $\theta = 0$ and $\theta = \pi$. The solutions are:

   $$\Theta(\theta) = \sqrt{\frac{2\ell+1}{2}\frac{(\ell-m)!}{(\ell+m)!}} P_\ell{}^m(\cos\theta) \tag{12.1.5}$$

   with $P_\ell{}^m(z)$ defined by eq. (10.6.10) or equivalently by eqs. (10.6.12) and (10.6.13). ℓ is a non-negative integer that can take any value from $|m|$ to infinity and $\beta = -\ell(\ell+1)$. This is also a Sturm-Liouville-type equation with eigenvalue $\lambda_\ell = \beta = -\ell(\ell+1)$. The inner product appropriate to the space of functions satisfying these boundary conditions is:

   $$(f, g) = \int_0^{\pi} f^*(\theta)g(\theta)\sin\theta\, d\theta \tag{12.1.6}$$

   and the eigenfunctions (12.1.5) have been chosen orthonormal with respect to this inner product.

3. $R(r)$:

   The radial equation is:

$$R''(r) + \frac{2}{r}R'(r) - \frac{\ell(\ell+1)}{r^2}R(r) = 0 \tag{12.1.7}$$

   with the boundary condition that $R(r)$ be bounded as $r \to 0$. Notice that the boundary condition on φ at $r = a$ is *inhomogeneous* and cannot be separately imposed on $R(r)$. It must therefore be held in reserve and not imposed until the functions $R(r)$, $\Theta(\theta)$ and $\Phi(\phi)$ are combined into solutions $\varphi(r, \theta, \phi)$ of the P.D.E..

   This is an Euler equation, *i.e.* an equation with precisely two regular singular points. The general solution to this equation is given in section 9.2:

$$R(r) = \gamma_1 r^\ell + \gamma_2 r^{-\ell-1} \tag{12.1.8}$$

   in which γ_1 and γ_2 are integration constants. The boundary condition at $r = 0$ implies that $\gamma_2 = 0$. It is important to realize that eq. (12.1.7) with only the boundary condition at $r = 0$ is *not* a Sturm-Liouville-type eigenvalue problem. (What would play the role of the eigenvalue?)

It is convenient to combine the angular solutions into the spherical harmonics, $Y_{\ell m}(\theta, \phi)$, introduced in eq. (11.2.19a). The separated solutions become:

$$\varphi_{\ell m}(r, \theta, \phi) = C r^\ell Y_{\ell m}(\theta, \phi). \tag{12.1.9}$$

C is an arbitrary constant.

Here comes the main argument. Since the solutions to the angular equations are the eigenfunctions of self-adjoint differential operators, they are known to be complete. This implies that *any* function $f(r, \theta, \phi)$ that obeys the θ and ϕ boundary conditions, regardless of whether or not it satisfies the differential equation (12.1.1) can be expanded in terms of these eigenfunctions:

$$f(r, \theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} k_{\ell m}(r) Y_{\ell m}(\theta, \phi) \tag{12.1.10}$$

The coefficients, $k_{\ell m}(r)$, in this expansion are written as functions of r because they may differ for differing values of r. Their r-dependence is determined, however, by the requirement that $f(r, \theta, \phi)$ satisfy the differential equation (12.1.1). This equation not surprisingly just implies that the functions $k_{\ell m}(r)$ must satisfy the radial equation (12.1.7). The solution regular at $r = 0$ is precisely (12.1.8): $k_{\ell m}(r) = c_{\ell m} r^\ell$.

The upshot is that we have proven the following result:

Theorem 12.1. *The most general (piecewise smooth etc.) solution to the homogeneous boundary-value problem given by the above P.D.E. subject to all of the boundary conditions except the inhomogeneous information at $r = a$ is:*

$$\varphi_{hom}(r, \theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell m} r^{\ell} Y_{\ell m}(\theta, \phi) \tag{12.1.11}$$

The problem remains to find which particular element of this vector space of solutions satisfies the inhomogeneous boundary condition at $r = a$. To do so requires solving the following equation:

$$V(\theta, \phi) = \varphi_{hom}(a, \theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell m} a^{\ell} Y_{\ell m}(\theta, \phi) \tag{12.1.12}$$

for the unknown coefficients $c_{\ell m}$. Happily enough, the solution is given explictly by eq. (11.2.9) which for this example reads:

$$a^{\ell} c_{\ell m} = (Y_{\ell m}, V) = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta [Y_{\ell m}(\theta, \phi)]^* V(\theta, \phi). \tag{12.1.13}$$

To make this perfectly explicit suppose that the given potential on the boundary at $r = a$ is:

$$V(\theta, \phi) = \begin{cases} +V_0 & \text{for } 0 \le \theta < \frac{\pi}{2} \\ -V_0 & \text{for } \frac{\pi}{2} \le \theta \le \pi \end{cases} \tag{12.1.14}$$

so (12.1.13) reduces to:

$$c_{\ell m} = \frac{V_0}{a^{\ell}} \sqrt{\frac{2\ell + 1}{2}} \delta_{m0} \int_0^1 [P_{\ell}(x) - P_{\ell}(-x)] \, dx \tag{12.1.15}$$

and the problem reduces to the evaluation of this integral over Legendre polynomials.

To evaluate this notice the following property satisfied by the Legendre polynomials: $P_{\ell}(-x) = (-)^{\ell} P_{\ell}(x)$. This is most easily proven by noticing that the generating function $g(x, t)$ defined by eq. (10.6.23) satisfies $g(-x, t) = g(x, -t)$. Inspection of eq. (12.1.15) shows that this implies that the coefficients $c_{\ell m}$ vanish if ℓ is even.

For $\ell = 2n + 1$ on the other hand the integral $\int_0^1 P_{2n+1}(x) \, dx$ may be evaluated for all n by directly integrating $g(x, t)$ with respect to x:

$$\begin{aligned} \sum_{\ell=0}^{\infty} t^{\ell} \int_0^1 P_{\ell}(x) \, dx &= \int_0^1 g(x, t) \, dx \\ &= \int_0^1 \frac{dx}{\sqrt{1 - 2xt + t^2}} \\ &= \frac{1}{t} \left[\sqrt{1 + t^2} - 1 + t \right] \\ &= 1 + \sum_{n=1}^{\infty} \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{3}{2} - n) n!} t^{2n-1}. \end{aligned} \tag{12.1.16}$$

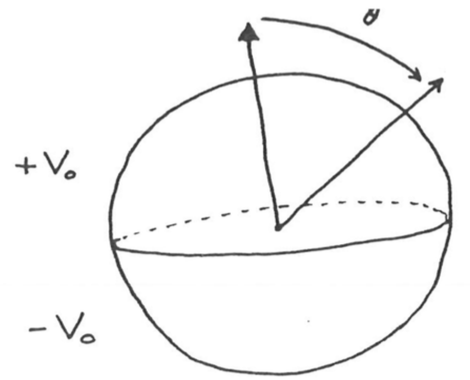

Figure 12.1. The Potential V_0

Reading off the coefficient of t^{2n+1} then gives:

$$\int_0^1 P_{2n+1}(x)\, dx = \frac{(-)^n(2n-1)!!}{2^{n+1}(n+1)!} \tag{12.1.17}$$

The only nonvanishing constants are therefore given by:

$$c_{2n+1,0} = \frac{V_0}{a^\ell}\sqrt{\frac{4n+3}{2}}\frac{(-)^n(2n-1)!!}{2^n(n+1)!} \tag{12.1.18}$$

and so the solution to the given boundary-value problem with boundary potential given by eq. (12.1.14) therefore is:

$$\varphi(r,\theta,\phi) = V_0\sum_{n=0}^{\infty}\sqrt{\frac{4n+3}{2}}\frac{(-)^n(2n-1)!!}{2^n(n+1)!}\left(\frac{r}{a}\right)^{2n+1} P_{2n+1}(\cos\theta). \tag{12.1.19}$$

12.2 Example II: The Diffusion Equation

Next consider an example using a parabolic differential equation, for which both boundary information and initial conditions are required.

Problem 12.2. *Consider the diffusion of particles within a cylindrical cavity of length L and radius a. The density of particles, $n(\rho,\xi,z,t)$, is specified at $t=0$ to be $n_0(\rho,\xi,z)$. No particles diffuse through the boundaries, so $\mathbf{n}\cdot\nabla n = 0$. Find the particle distribution at later times t.*

The differential equation governing diffusion is, in cylindrical coordinates, eq. (5.2.3):

$$\frac{\partial n}{\partial t} - \kappa\left[\frac{\partial^2 n}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial n}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 n}{\partial \xi^2} + \frac{\partial^2 n}{\partial z^2}\right] = 0. \tag{12.2.1}$$

and the boundary conditions are:

$$\left.\frac{\partial n}{\partial z}\right|_{z=0} = \left.\frac{\partial n}{\partial z}\right|_{z=L} = 0, \tag{12.2.2}$$

$$\left.\frac{\partial n}{\partial \rho}\right|_{\rho=a} = 0, \tag{12.2.3}$$

and the initial condition is

$$n(\rho, \xi, z, 0) = n_0(\rho, \xi, z). \tag{12.2.4}$$

The first step is to generate the separated solutions $n(\rho, \xi, z, t) = R(\rho)\Xi(\xi)Z(z)T(t)$ by separating variables in cylindrical coordinates. The separated equations together with their boundary conditions are:

1. $T(t)$:

   The O.D.E. is:

   $$T'(t) = \alpha T(t). \tag{12.2.5}$$

   The temporal boundary condition (12.2.4) *cannot* be imposed solely on the separated function $T(t)$. The solution to this equation are, up to an overall multiplicative constant:

   $$T(t) = \exp(\alpha t). \tag{12.2.6}$$

   This is *not* a Sturm-Liouville eigenvalue problem.

2. $\Xi(\xi)$:

   This coordinate satisfies the Sturm-Liouville problem:

   $$\Xi''(\xi) = \beta \Xi(\xi) \tag{12.2.7}$$

   with $\Xi(\xi + 2\pi) = \Xi(\xi)$. This has eigenvalues $\beta = -m^2$, for integer m, normalized eigenfunctions:

   $$\Xi_m(\xi) = \frac{1}{\sqrt{2\pi}} e^{im\xi}, \tag{12.2.8}$$

   and inner product:

   $$(f, g) = \int_0^{2\pi} f^*(\xi) g(\xi)\, d\xi \tag{12.2.9}$$

3. $Z(z)$:

   This coordinate satisfies the Sturm-Liouville problem:

   $$Z''(z) = \gamma Z(z) \tag{12.2.10}$$

with $Z'(0) = Z'(L) = 0$. This has eigenvalues $\gamma = -(\pi n/L)^2$, for non-negative integer n, normalized eigenfunctions:

$$Z_n(z) = \sqrt{\frac{2}{L}} \cos\left(\frac{\pi n z}{L}\right), \tag{12.2.11}$$

(except for $n = 0$ for which $Z_0 = 1/\sqrt{L}$) and inner product:

$$(f, g) = \int_0^L f^*(z) g(z)\, dz \tag{12.2.12}$$

4. $R(r)$:

   The radial equation:

$$R''(\rho) + \frac{1}{\rho} R'(\rho) - \left[\left(\frac{\pi^2 n^2}{L^2} + \frac{\alpha}{\kappa}\right) + \frac{m^2}{\rho^2}\right] R(\rho) = 0 \tag{12.2.13}$$

   is Bessel's equation. The boundary conditions are that $R(\rho)$ be bounded as $\rho \to 0$ and that $R'(a) = 0$. These together with the O.D.E. (12.2.13) form a Sturm-Liouville problem with eigenvalue $\lambda = \alpha/\kappa$. The solution that is regular at the origin is the Bessel function $J_m(\zeta\rho)$ if $-\zeta^2 = (\alpha/\kappa + \pi^2 n^2/L^2)$ is negative. On the other hand,if this quantity is positive then the regular solution is the modified Bessel function $I_m(\tilde{\zeta}\rho)$ with $(\tilde{\zeta})^2 = (\alpha/\kappa + \pi^2 n^2/L^2)$. To decide which of these is the correct case impose the boundary condition at $\rho = a$: $R'(a) = 0$.

Claim: The equation $I'_m(\tilde{\zeta} a) = 0$ has no solutions for real $\tilde{\zeta} a$.

Proof. To prove this use the following argument. From the integral representation

$$I_m(x) = \frac{1}{\Gamma(m + \frac{1}{2})\sqrt{\pi}} \int_{-1}^{1} ds\, e^{-xs} (1 - s^2)^{m - \frac{1}{2}} \tag{12.2.14}$$

it is clear that $I_m(x) > 0$ for all $x > 0$. Furthermore, the recursion relation:

$$I'_m(x) = \frac{1}{2}[I_m(x) + I_{m+1}(x)] \tag{12.2.15}$$

then implies the same for $I'_m(x)$. This conclusion is not true for $J_m(x)$ however as may be seen from its asymptotic form (10.7.19). From eq. (10.7.19) we see that there are an infinite number of real zeroes to both $J_m(x)$ and $J'_m(x)$. □

The boundary condition at $\rho = a$ therefore implies that $R_{mnl}(\rho) = C J_m(\zeta_{ml}\rho)$ where ζ_{ml} with $l = 1, 2, ...$ are the roots to $J'_m(\zeta_{ml} a) = 0$. This in particular implies that the constant, α, is determined to be α_{mnl} given by the combination $\alpha_{mnl} = -\kappa(\pi^2 n^2/L^2 + \zeta^2_{ml})$ (which is negative).

The eigenvalues are therefore

$$\lambda_{mnl} = \frac{\alpha_{mnl}}{\kappa} = -\left(\frac{\pi^2 n^2}{L^2} + \zeta_{ml}^2\right) \tag{12.2.16}$$

and the corresponding eigenfunctions are

$$R_{ml}(\rho) = N_{ml} J_m(\zeta_{ml}\rho) \tag{12.2.17}$$

with normalization constant, N_{ml}, chosen to normalize the eigenfunction. Since the inner product in this case is:

$$(f, g) = \int_0^a f^*(\rho) g(\rho) \rho \, d\rho \tag{12.2.18}$$

the normalization constant works out to be:

$$\begin{aligned} \frac{1}{N_{ml}^2} &= \int_0^a |J_m(\zeta_{ml}\rho)|^2 \rho \, d\rho \\ &= \left(\frac{a^2}{2} - \frac{m^2}{2\zeta_{ml}^2}\right) [J_m(\zeta_{ml} a)]^2 \end{aligned} \tag{12.2.19}$$

The logic of the argument now proceeds as for the previous example. Define the separated eigenfunctions:

$$\psi_{mnl}(\rho, \xi, z) = \sqrt{\frac{1}{\pi L}} N_{ml} J_m(\zeta_{ml}\rho) \cos\left(\frac{\pi n z}{L}\right) e^{im\xi} \tag{12.2.20}$$

The completeness of the eigenfunctions $R_{ml}(\rho)$, $\Xi_m(\xi)$ and $Z_n(z)$ ensures that *any* function satisfying the ξ, z and $\rho = 0$ boundary conditions can be expanded in terms of the $\psi_{mnl}(\rho, \xi, z)$'s:

$$f(\rho, \xi, z, t) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{l=1}^{\infty} k_{mnl}(t) \psi_{mnl}(\rho, \xi, z) \tag{12.2.21}$$

regardless of whether it satisfies the diffusion equation (12.2.1). The extra information that the function satisfies this P.D.E. determines the t-dependence of the coefficients to be $k_{mnl}(t) = c_{mnl} \exp(\alpha_{mnl} t)$ and so an arbitrary solution to the homogeneous part of the original boundary-value problem is:

$$n_{hom}(\rho, \xi, z, t) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{l=1}^{\infty} c_{mnl} e^{\alpha_{mnl} t} \psi_{mnl}(\rho, \xi, z) \tag{12.2.22}$$

The specific solution satisfying the initial condition (12.2.4) is found by inverting

$$n_0(\rho, \xi, z) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{l=1}^{\infty} c_{mnl} \psi_{mnl}(\rho, \xi, z) \tag{12.2.23}$$

for the coefficients c_{mnl} using relations (11.2.9). The result is:

$$\begin{aligned} c_{mnl} &= (\psi_{mnl}, n_0) \\ &= \int_0^{2\pi} d\xi \int_0^L dz \int_0^a d\rho\, \rho\psi^*_{mnl}(\rho, \xi, z) n_0(\rho, \xi, z). \end{aligned} \tag{12.2.24}$$

For example suppose the initial particle distribution is given by the following:

$$n_0 = \begin{cases} 0 & \text{if } \rho > \epsilon \\ \eta & \text{if } 0 \le \rho \le \epsilon\,. \end{cases} \tag{12.2.25}$$

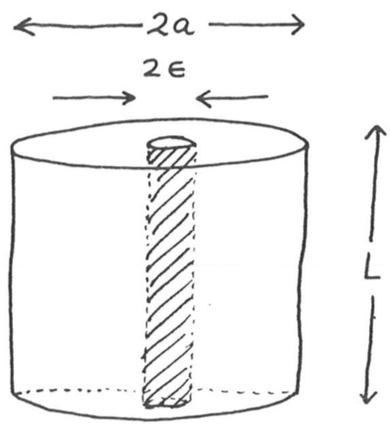

Figure 12.2. The Initial Particle Distribution n_0

Then eq. (12.2.24) becomes:

$$c_{mnl} = \eta \frac{2\sqrt{\pi L}}{aJ_0(\zeta_{0l}a)} \delta_{m0}\delta_{n0} \int_0^\epsilon \rho J_0(\zeta_{0l}\rho)\, d\rho \tag{12.2.26}$$

The recursion relation $J_1(x) = x[J_0(x) + J_1'(x)]$ may be used to evaluate the Bessel function integrals and so to get:

$$c_{mnl} = 2\eta\epsilon \frac{\sqrt{\pi L} J_1(\zeta_{0l}\epsilon)}{a\zeta_{0l} J_0(\zeta_{0l}a)} \delta_{m0}\delta_{n0} \tag{12.2.27}$$

The solution to the initial-value problem is:

$$n(\rho, \xi, z, t) = \frac{2\eta\epsilon}{a^2} \sum_{l=1}^{\infty} \frac{J_1(\zeta_{0l}\epsilon)}{\zeta_{0l}[J_0(\zeta_{0l}a)]^2} e^{-\kappa\zeta_{0l}^2 t} J_0(\zeta_{0l}\rho). \tag{12.2.28}$$

This completes our discussion of boundary-value problems.