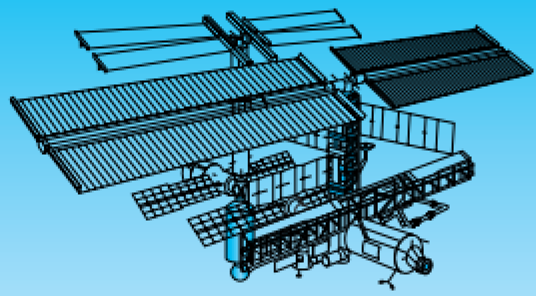
CHAPTER 8



Series Solutions of Differential Equations, Special Functions, and Sturm–Liouville Equations

Linear second order variable coefficient equations arise in many applications, but only in a few special cases is it possible to express their general solution of a finite linear combination of elementary functions. As analytical, rather than purely numerical, information about solutions is often essential, some other way must be found to represent the solutions of such equations. The approach developed in this chapter involves seeking solutions of certain types of equation in the form of power series, and in other cases using an approach due to Frobenius that involves seeking solutions in the form of power series multiplied by a factor x^c, where c is not an integer. Applications are made to a number of typical linear variable coefficient equations, and then to the important Legendre, Chebyshev, and Bessel equations that lead in turn to Legendre and Chebyshev polynomials and to Bessel functions.

Two-point boundary value problems, called Sturm–Liouville systems, that are defined over an interval $a \le x \le b$ and contain a parameter λ are introduced. It is shown that their solutions only exist for an infinite number of special values of the parameter $\lambda_1, \lambda_2, \ldots,$ called the eigenvalues of the problem. Each solution $\varphi_n(x)$ corresponding to an eigenvalue λ_n is called an eigenfunction, and the eigenfunctions are shown to have the special property of orthogonality with respect to a function $w(x)$ called the weight function. This means that if the set of eigenfunctions is $\{\varphi_n(x)\}_{n=1}^{\infty}$, the integral $\int_a^b \varphi_m(x)\varphi_n(x)w(x)dx$ is positive when $n = m$ and zero when $n \neq m$. This property will be used extensively in Chapter 18 when solving partial differential equations.

Fundamental properties of eigenfunctions and eigenvalues are established for general Sturm–Liouville systems, after which a number of frequently occurring and important special cases are examined.

8.1 A First Approach to Power Series Solutions of Differential Equations

The solutions of many differential equations can be expressed in terms of elementary functions such as sine, cosine, exponential, and logarithm, all of whose mathematical properties are well known. When required, the analytical behavior of solutions that involve elementary functions can be explored by making use of their familiar properties. Numerical solutions are obtained easily, either by using a pocket calculator to find the values of the elementary functions involved, or through the use of standard subroutines that form a part of all basic mathematical software packages. With either a pocket calculator or a software package, the method of calculating functional values is usually based on a series expansion of the function concerned.

Most differential equations cannot be solved in terms of elementary functions, yet some form of analytical solution is often needed rather than a purely numerical one, so the fundamental question that then arises is how to obtain a solution in the form of a series, when only the differential equation is known. It is the purpose of this chapter to answer this question, and in the process to show how the form of series solution obtained depends on what are called the *singular points* of the differential equation.

We begin our approach to this problem by showing how series solutions can be found for first and second order linear differential equations with initial conditions specified at $x = x_0$. The series we obtain will be in powers of $x - x_0$, and they will be said to be *expanded* about the point x_0. The first order linear differential equation will be assumed to be of the form

$$y' + p(x)y = r(x) \quad \text{with } y(x_0) = y_0, \tag{1}$$

and the second order linear differential equation will be assumed to be of the form

$$y'' + P(x)y' + Q(x)y = R(x) \quad \text{with } y(x_0) = y_0,\ y'(x_0) = y_1, \tag{2}$$

where the functions $p(x)$, $r(x)$, $P(x)$, $Q(x)$, and $R(x)$ can all be expanded as Taylor series about the point x_0.

analytic in a neighborhood

Functions with this property are said to be **analytic** in a **neighborhood** of the point x_0 or, more simply, to be **analytic** at x_0. The method to be developed will be seen to be capable of extension to a higher order linear differential equation in an obvious manner, provided only that the coefficients of y and its derivatives that are involved and the nonhomogeneous term are analytic at x_0.

how to find a power series solution

The approach is best illustrated by considering equation (1), and seeking a solution about x_0 of the form

$$\begin{aligned} y(x) &= y(x_0) + (x - x_0)y'(x_0) + \frac{(x-x_0)^2}{2!}y''(x_0) + \frac{(x-x_0)^3}{3!}y'''(x_0) + \cdots \\ &= \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{n!} y^{(n)}(x_0), \quad \text{with } y^{(n)}(x) = d^n y/dx^n. \end{aligned} \tag{3}$$

Setting $x = x_0$ in (1) gives

$$y^{(1)}(x_0) + p(x_0)y(x_0) = r(x_0),$$

but $y(x_0) = y_0$, so

$$\begin{aligned} y^{(1)}(x_0) &= r(x_0) - p(x_0)y(x_0) \\ &= r(x_0) - p(x_0)y_0. \end{aligned}$$

To determine $y^{(2)}(x)$ we differentiate equation (1) once with respect to x to obtain

$$y^{(2)}(x) + p^{(1)}(x)y(x) + p(x)y^{(1)}(x) = r^{(1)}(x),$$

where $p^{(1)}(x) = p'(x)$ and $r^{(1)}(x) = r'(x)$. Then, after setting $x = x_0$ and using the fact that $y^{(1)}(x_0) = r(x_0) - p(x_0)y_0$, we find that

$$y^{(2)}(x_0) = r^{(1)}(x_0) - p^{(1)}(x_0)y_0 - p(x_0)[r(x_0) - p(x_0)y_0].$$

Higher order derivatives $y^{(n)}(x_0)$ can be computed in similar fashion by repeated differentiation of the original differential equation coupled with the use of lower order derivatives that have already been determined. Once the values of $y^{(k)}(x_0)$ have been found for $k = 1, 2, \ldots, N$, for some given integer N, substitution into series (3) provides the required approximation to the power series solution of the initial value problem for the differential equation up to terms of order $(x - x_0)^N$. The existence and uniqueness of the solution are guaranteed by Theorem 5.2.

This method generates the Taylor series expansion of $y(x)$ about the point x_0 when $x_0 \neq 0$, and its Maclaurin series expansion when $x_0 = 0$, though these series are often simply called power series about $x_0 \neq 0$ and $x_0 = 0$, respectively.

EXAMPLE 8.1 Find the first five terms in the series solution of

$$y' + (1 + x^2)y = \sin x, \quad \text{with } y(0) = a.$$

Solution As the initial condition is specified at $x = 0$, the power series solution is an expansion about the origin and so is, in fact, a Maclaurin series. The functions $1 + x^2$ and $\cos x$ are analytic for all x, so the series expansion can certainly be found about the origin.

Setting $x = 0$ in the equation and substituting for the initial conditions shows that $y'(0) = y^{(1)}(0) = -a$. Differentiation of the differential equation gives

$$y^{(2)} + 2xy + (1 + x^2)y^{(1)} = \cos x,$$

where $y^{(2)} = y''$, so setting $x = 0$ this becomes

$$y^{(2)}(0) + y^{(1)}(0) = 1,$$

but $y^{(1)}(0) = -a$ and so $y^{(2)}(0) = 1 + a$. Repeating this process to find higher order derivatives leads to the results $y^{(3)}(0) = -(1 + 3a)$, $y^{(4)}(0) = 9a, \ldots$. Substituting these results into series (3) shows that, to terms of order x^4, the required solution takes the form

$$y(x) = a - ax + (1 + a)\frac{x^2}{2!} - (1 + 3a)\frac{x^3}{3!} + 9a\frac{x^4}{4!} + \cdots.$$

EXAMPLE 8.2 Find the first five terms in the series solution of

$$y' + 4xy = 3e^{x-1}, \quad \text{with } y(1) = 1.$$

Solution In this case the functions x and e^{x-1} are analytic for all x, but as the expansion is about $x = 1$, the power series solution that is obtained will be a Taylor series expansion about the point $x = 1$. Setting $x = 1$ in the differential equation and using the initial condition $y(1) = 1$ shows that $y^{(1)}(1) = -1$.

Differentiation of the differential equation gives

$$y^{(2)} + 4y + 4xy^{(1)} = 3e^{x-1},$$

so setting $x = 1$ and using the result $y^{(1)}(1) = -1$ shows that $y^{(2)}(1) = 3$.

Repeating this process leads to the results that $y^{(3)}(1) = -1$ and $y^{(4)}(1) = -29$, so substituting into (3) shows that the Taylor series expansion of the solution up to terms of order $(x-1)^4$ is

$$y(x) = 1 - (x-1) + \frac{3}{2}(x-1)^2 - \frac{1}{6}(x-1)^3 - \frac{29}{24}(x-1)^4 + \cdots. \quad \blacksquare$$

This same method can be applied to a second order equation of the type shown in (2), though a more general approach will be developed later to deal with the case in which the first term is of the form $a(x)y''(x)$, and the expansion is about a point x_0 where $a(x_0) = 0$.

EXAMPLE 8.3 Find the terms up to x^5 in the series solution of

$$y'' + xy' + (1 - x^2)y = x \quad \text{with } y(0) = a,\ y'(0) = b.$$

Solution The coefficients x and $(1 - x^2)$ and the nonhomogeneous term x are analytic for all x, so as the initial data is given at $x = 0$, a Maclaurin series solution can be found.

Setting $x = 0$ in the equation and using the initial conditions $y(0) = a$ and $y'(0) = b$ gives $y^{(2)}(0) = -a$. Differentiating the differential equation we have

$$y^{(3)} + y^{(1)} + xy^{(2)} - 2xy + (1 - x^2)y^{(1)} = 1,$$

so setting $x = 0$ and using the results $y^{(2)}(0) = -a$ and $y^{(1)}(0) = b$ shows that $y^{(3)}(0) = 1 - 2b$. A repetition of this process leads to the results $y^{(4)}(0) = 5a$, $y^{(5)}(0) = 14b - 4, \ldots$, so substituting into (3) shows that to terms of order x^5 the Maclaurin series expansion of the solution is

$$y(x) = a + bx - \frac{1}{2}ax^2 + \left(\frac{1-2b}{6}\right)x^3 + \frac{5a}{24}x^4 + \left(\frac{7b-2}{60}\right)x^5 + \cdots. \quad \blacksquare$$

Summary

Often a variable coefficient equation cannot be solved in terms of known functions, though some form of analytical solution is still required. This section has shown how to overcome this difficulty in some cases by finding a solution in terms of a power series expanded about a point of interest $x = a$. The method was seen to work provided the functions in the equation have Taylor series expansions about $x = a$. It will be shown later how to find series solutions in a systematic manner, and also how to generalize this approach to other types of equation.

EXERCISES 8.1

Find the first five terms in the power series solution of the following initial value problems.

1. $y' + (1 + x^2)y = x^2$, with $y(0) = 1$.
2. $2y' + xy = 1 - x$, with $y(0) = 2$.
3. $y' + (1 - 2x)y = x$, with $y(0) = -1$.
4. $4y' + (1 + x + x^2)y = x$, with $y(0) = 3$.
5. $y' + (x - 2x^2)y = 1$, with $y(0) = 1$.
6. $y' - 2xy = 1 - x$, with $y(0) = 2$.
7. $3y' + (1 - x^2)y = 1$, with $y(0) = 2$.
8. $y' + (1 + x)y = 1 + x^2$, with $y(0) = 1$.
9. $y'' - 2xy' + x^2y = 0$, with $y(0) = a$, $y'(0) = b$.
10. $2y'' + 2(1 + x)y' - y = 0$, with $y(0) = a$, $y'(0) = b$.

11. $(1+x^2)y'' + 3xy' + (1-x^2)y = 1 + x$, with $y(0) = a$, $y'(0) = b$.
12. $(1+3x^2)y'' + 2xy' + 2xy = 1$, with $y(0) = a$, $y'(0) = b$.
13. $y'' + 7y' + x^2y = 0$, with $y(0) = a$, $y'(0) = b$.
14. $xy'' + (1+x)y' + xy = b$, with $y(0) = a$, $y'(0) = 0$.
15. $2y'' + 3x^2y' + (1-x^2)y = 2x$, with $y(0) = a$, $y'(0) = b$.
16. $3y'' + 2xy' + (1-2x^2)y = 1 + 2x$, with $y(0) = a$, $y'(0) = b$.

8.2 A General Approach to Power Series Solutions of Homogeneous Equations

The method developed in Section 8.1 works satisfactorily if only the first few terms in a power series solution are required, but it has the disadvantage that a separate calculation is required each time a coefficient is determined. The present section shows how in many cases this difficulty can be overcome by introducing a systematic and simple way of generating arbitrarily many terms in a power series solution of the homogeneous linear differential equation

$$a(x)y'' + b(x)y' + c(x)y = 0 \tag{4}$$

about a point x_0, when $a(x)$, $b(x)$, and $c(x)$ are polynomials with $a(x_0) \neq 0$.

The approach enables the coefficients of the power series solution to be determined by means of a *recurrence relation* that relates a few consecutive coefficients in the series. This has the advantage that once the first few coefficients in the series expansion have been found, the rest can be generated by means of the recurrence relation.

There will be no loss of generality if the approach is based on an expansion about the origin, because if one is required about an arbitrary point $x = x_0$, the change of variable $X = x - x_0$ will shift the point $x = x_0$ to $X = 0$. For example, suppose a solution of

$$y'' + (2 + 3x)y' + x^2y = 0$$

is required about the point $x = 1$, corresponding to the specification of the initial conditions for $y(1)$ and $y'(1)$ at $x = 1$. Setting $X = x - 1$ and $y(x) = Y(x-1) = Y(X)$, it follows that $y(1) = Y(0)$, $dy/dx = dY/dX$, $d^2y/dx^2 = d^2Y/dX^2$, and $x = X + 1$, so in terms of the new variables X and Y the equation and initial conditions become

$$Y'' + (5 + 3X)Y' + (1 + X)^2Y = 0, \quad \text{with } Y(0) = y(1),\ Y'(0) = y'(1).$$

Setting $X = x - 1$ in the power series solution of this equation expanded about $X = 0$ reduces it to the solution of the original equation expanded about $x = 1$.

The approach we now describe involves seeking a solution in the form of a general power series

$$y(x) = \sum_{n=0}^{\infty} a_n x^n \tag{5}$$

and finding a relationship between the coefficients a_n by substituting (5) into the homogeneous differential equation

$$a(x)y'' + b(x)y' + c(x)y = 0. \tag{6}$$

We will assume that the coefficients $a(x)$, $b(x)$, and $c(x)$ in the differential equation are polynomials in x, and so are analytic at $x = 0$, and also that $a(0) \neq 0$. If (5) is to be a solution of (6), it must satisfy the differential equation for all x, but this will only be possible if, after combining terms, the coefficient of each power of x in the new power series is zero. It will be seen later that it is this last requirement that leads to the determination of the coefficients a_n in terms of a recurrence relation.

Before illustrating the approach by means of an example, we first find expressions for the derivatives $y'(x)$ and $y''(x)$ that will be needed in the calculation. Writing out the first few terms of $y(x)$ in (5) gives

$$y(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots = \sum_{n=0}^{\infty} a_n x^n. \tag{7}$$

Differentiating this expression term by term with respect to x, which is permitted for x inside the interval of convergence of the series, we arrive at the result

$$y'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \cdots = \sum_{n=1}^{\infty} n a_n x^{n-1}, \tag{8}$$

and after a further differentiation we have

$$y''(x) = 2a_2 + 2 \cdot 3a_3 x + 3 \cdot 4a_4 x^2 + \cdots = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}. \tag{9}$$

In what is to follow it will be important to remember that the summation in (8) starts at $n = 1$, whereas the summation in (9) starts at $n = 2$.

EXAMPLE 8.4 Find the recurrence relation that must be satisfied by coefficients in the series solution of the differential equation

$$y'' + 2xy' + (1 + x^2)y = 0$$

when the expansion is about the origin. Solve the initial value problem for this differential equation given that $y(0) = 3$ and $y'(0) = -1$.

Solution Substituting $y(x) = \sum_{n=0}^{\infty} a_n x^n$ into the differential equation and using (8) and (9) gives

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} + 2x \sum_{n=1}^{\infty} n a_n x^{n-1} + (1 + x^2) \sum_{n=0}^{\infty} a_n x^n = 0.$$

Taking the factor $2x$ in the second term and the factor x^2 in the third term under their respective summation signs allows the equation to be written in the form

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} + \sum_{n=1}^{\infty} 2na_n x^n + \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} a_n x^{n+2} = 0.$$

The powers of x in the first and last summations are different from those in the middle two summations, so before combining the summations in order to find the coefficient of each power of x, it will first be necessary to change the power of x in the first and last terms from $n-2$ and $n+2$ to n.

In the first summation we set $m = n - 2$, causing the summation to become

$$\sum_{m=0}^{\infty} (m+2)(m+1)a_{m+2}x^m.$$

However, m is simply a summation index that can be replaced by any other symbol, so we will replace it by n to obtain the equivalent expression

$$\sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2}x^n.$$

Similarly, by setting $m = n + 2$ in the last summation, and then replacing m by n, we find that

$$\sum_{n=0}^{\infty} a_n x^{n+2} \quad \text{becomes} \quad \sum_{n=2}^{\infty} a_{n-2}x^n.$$

We now substitute these last two results into the series solution of the differential equation to obtain

$$\sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2}x^n + \sum_{n=1}^{\infty} 2na_n x^n + \sum_{n=0}^{\infty} a_n x^n + \sum_{n=2}^{\infty} a_{n-2}x^n = 0,$$

where now each summation involves x^n, though not all summations start from $n = 0$.

Separating out the terms corresponding to $n = 0$ and $n = 1$, and collecting all the remaining terms under a single summation sign in which the summation starts from $n = 2$, this becomes

$$2a_2 + a_0 + (6a_3 + 3a_1)x + \sum_{n=2}^{\infty} [(n+2)(n+1)a_{n+2} + 3a_n + a_{n-2}]x^n = 0.$$

As already remarked, if this power series is to be a solution of the differential equation it must satisfy the equation identically for all x, but this will only be possible if in the foregoing expression the coefficient of each power of x vanishes. Applying this condition to the preceding series we find that for it to vanish identically for all x,

deriving and using a recurrence relation

$$\text{(coefficient of } x^0) \quad 2a_2 + a_0 = 0$$
$$\text{(coefficient of } x) \quad 6a_3 + 3a_1 = 0$$

and

$$\text{(coefficient of } x^n) \quad (n+2)(n+1)a_{n+2} + 3a_n + a_{n-2} = 0, \quad \text{for } n \geq 2.$$

The first condition shows that

$$a_2 = -\frac{1}{2}a_0,$$

while the second condition shows that

$$a_3 = -\frac{1}{2}a_1,$$

where a_0 and a_1 are arbitrary constants.

The third condition is a **recurrence relation** (also called a **recursion relation** or an **algorithm**) that in this case relates three coefficients whose indices differ by 2, so given a_{n-2} and a_n we can find a_{n+2} for $n = 2, 3, 4, \ldots$.

We now show how to determine the first few coefficients a_n by writing the recursion relation in the form

$$a_{n+2} = -\frac{[(2n+1)a_n + a_{n-2}]}{(n+1)(n+2)}$$

and setting $n = 2, 3, 4, \ldots$.

For $n = 2$, after using $a_2 = -\frac{1}{2}a_0$, we find that

$$a_4 = -\frac{(5a_2 + a_0)}{12} = \frac{a_0}{8},$$

whereas for $n = 3$, after using $a_3 = -\frac{1}{2}a_1$, we find that

$$a_5 = -\frac{(7a_3 + a_1)}{20} = \frac{a_1}{8}.$$

Continuing this process generates the coefficients

$$a_6 = -\frac{a_0}{48}, \quad a_7 = -\frac{a_1}{48}, \quad a_8 = \frac{a_0}{384}, \quad a_9 = \frac{a_1}{384}, \ldots.$$

Thus, all the coefficients with even suffixes are determined in terms of the arbitrary constant a_0, whereas all the coefficients with odd suffixes are determined in terms of the arbitrary constant a_1.

Substituting these coefficients into the power series $y(x) = \sum_{n=0}^{\infty} a_n x^n$ and grouping terms gives

$$y(x) = a_0\left(1 - \frac{1}{2}x^2 + \frac{1}{8}x^4 - \frac{1}{48}x^6 + \frac{1}{384}x^8 - \cdots\right)$$
$$+ a_1\left(x - \frac{1}{2}x^3 + \frac{1}{8}x^5 - \frac{1}{48}x^7 + \frac{1}{384}x^9 - \cdots\right).$$

As the coefficients a_0 and a_1 are arbitrary, the functions represented by the series

$$y_1(x) = 1 - \frac{1}{2}x^2 + \frac{1}{8}x^4 - \frac{1}{48}x^6 + \frac{1}{384}x^8 - \cdots$$

and

$$y_2(x) = x - \frac{1}{2}x^3 + \frac{1}{8}x^5 - \frac{1}{48}x^7 + \frac{1}{384}x^9 - \cdots$$

are seen to be the two linearly independent solutions known to be associated with a homogeneous linear second order equation. So all possible solutions of the differential equation can be written in the form

$$y(x) = C_1 y_1(x) + C_2 y_2(x),$$

with C_1 and C_2 arbitrary constants, where to reconcile this result with our previous notation we notice that C_1 and C_2 have been written in place of a_0 and a_1.

To solve the initial value problem the constants C_1 and C_2 must be chosen such that $y(0) = 3$ and $y'(0) = -1$, so

$$3 = C_1 y_1(0) + C_2 y_2(0) \quad \text{and} \quad -1 = C_1 y_1'(0) + C_2 y_2'(0),$$

but $y_1(0) = 1$, $y_2(0) = 0$, and differentiation of the expressions for $y_1(x)$ and $y_2(x)$ shows that $y_1'(0) = 0$ and $y_2'(0) = 1$, so solving for C_1 and C_2 gives $C_1 = 3$ and $C_2 = -1$, showing that the required solution to the initial value problem is

$$y(x) = 3y_1(x) - y_2(x).$$

The coefficients of the power series expansions for $y_1(x)$ and $y_2(x)$ in the last example were sufficiently complicated that no attempt was made to deduce their general forms and they were merely generated from the recurrence relation. The next example is simpler, and we use it to illustrate the type of argument that is necessary when attempting to arrive at the form of the general term in a power series solution of a homogeneous linear differential equation. There are no specific rules to follow when seeking the form of a general term in a series, and success depends on experience and the ability to recognize the pattern of signs and numbers forming the coefficients.

EXAMPLE 8.5

Find two linearly independent solutions of

$$y'' + xy' + y = 0,$$

when the series expansion is about the origin, and hence solve the initial value problem for which $y(0) = 1$ and $y'(0) = 0$.

Solution Substituting results (7) to (9) into the differential equation gives

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} + x\sum_{n=1}^{\infty} na_n x^{n-1} + \sum_{n=0}^{\infty} a_n x^n = 0.$$

Shifting the summation index in the first term, taking the factor x under the second summation and separating out the constant term, as in Example 8.4, gives

$$2a_2 + a_0 + \sum_{n=1}^{\infty}[(n+2)(n+1)a_{n+2} + (n+1)a_n]x^n = 0.$$

Equating the coefficient of each power of x to zero, as in Example 8.4, shows that

$$2a_2 + a_0 = 0, \quad \text{so } a_2 = -\frac{a_0}{2},$$

and

$$(n+2)(n+1)a_{n+2} + (n+1)a_n = 0 \quad \text{for } n \geq 1,$$

but as $n + 1 \neq 0$ this last condition reduces to the simpler *recurrence relation*

$$a_{n+2} = -\frac{a_n}{n+2}, \quad \text{for } n = 1, 2, \ldots .$$

It follows directly from the recurrence relation that all even coefficients are multiples of a_0 and all odd coefficients are multiples of a_1 with

$$a_3 = -\frac{a_1}{3}, \quad a_4 = -\frac{a_2}{4} = \frac{a_0}{2\cdot 4}, \quad a_5 = -\frac{a_3}{5} = \frac{a_1}{3\cdot 5}, \quad a_6 = -\frac{a_4}{6} = -\frac{a_0}{2\cdot 4\cdot 6},$$

$$a_7 = -\frac{a_5}{7} = -\frac{a_1}{3\cdot 5\cdot 7}, \quad a_8 = -\frac{a_6}{8} = \frac{a_0}{2\cdot 4\cdot 6\cdot 8}, \quad a_9 = -\frac{a_7}{9} = \frac{a_1}{3\cdot 5\cdot 7\cdot 9}, \ldots,$$

where a_0 and a_1 are arbitrary constants.

It is apparent that the pattern of coefficients with even suffixes differs from the one for coefficients with odd suffixes, so each must be considered separately. Starting with the coefficients with even suffixes, we use the fact that if $m = 1, 2, \ldots,$ then $2m$ is an even number. A little experimentation shows that the signs of the terms with even suffixes are given by the factor $(-1)^m$.

Noticing that a_2, a_4, a_6, and a_8 can be written in the form

$$a_2 = \frac{(-1)a_0}{2}, \quad a_4 = \frac{1}{2\cdot 4}\frac{(-1)^2 a_0}{2^2 2!}, \quad a_6 = \frac{-a_0}{2\cdot 4\cdot 6} = \frac{(-1)^3 a_0}{2^3 3!},$$

$$a_8 = \frac{(-1)^4 a_0}{2\cdot 4\cdot 6\cdot 8} = \frac{(-1)^4 a_0}{2^4 4!}$$

suggests that if we set $n = 2m$, for $m = 0, 1, 2, \ldots,$ the even numbered terms can be written

$$a_{2m} = \frac{(-1)^m}{2^m m!}a_0.$$

A formal proof that this is the general coefficient in the series involving even powers of x can be obtained by mathematical induction, but we leave this as an exercise.

It is now necessary to consider the coefficients with odd suffixes, and to do this we use the fact that if $m = 1, 2, 3, \ldots,$ then $2m + 1$ is an odd number. Noticing that the coefficients a_3, a_5, a_7, and a_9 can be written

$$a_3 = \frac{-a_1}{3} = \frac{(-1)2a_1}{3!}, \quad a_5 = \frac{a_1}{3\cdot 5} = \frac{(-1)^2\, 2\cdot 4a_1}{1\cdot 2\cdot 3\cdot 4\cdot 5} = \frac{(-1)^2\, 2^2 2!}{5!},$$

$$a_7 = \frac{-a_1}{3\cdot 5\cdot 7} = \frac{(-1)^3\, 2\cdot 4\cdot 6a_1}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7} = \frac{(-1)^3\, 2^3 3! a_1}{7!},$$

$$a_9 = \frac{a_1}{3\cdot 5\cdot 7\cdot 9} = \frac{(-1)^4\, 2\cdot 4\cdot 6\cdot 8a_1}{9!} = \frac{(-1)^4\, 2^4 4! a_1}{9!}$$

suggests that the coefficients in the series of odd powers of x can be written

$$a_{2m+1} = \frac{(-1)^m\, 2^m m!}{(2m+1)!}a_1.$$

Here again we leave as an exercise the task of giving an inductive proof that this is, indeed, the coefficient of the general term in the series involving odd powers of x.

The solution of the differential equation has now separated into two series, one multiplied by a_0 containing only even powers of x and the other multiplied by a_1 containing only odd powers of x, so the solution becomes

$$y(x) = a_0 \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^m m!} + a_1 \sum_{m=0}^{\infty} \frac{(-1)^m\, 2^m m! x^{2m+1}}{(2m+1)!}.$$

As a_0 and a_1 are arbitrary constants, and the two series are not proportional, it follows that two linearly independent solutions of the differential equation are

$$y_1(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^m m!} \quad \text{and} \quad y_2(x) = \sum_{m=0}^{\infty} \frac{(-1)^m 2^m m! x^{2m+1}}{(2m+1)!},$$

so the general solution is

$$y(x) = C_1 y_1(x) + C_2 y_2(x),$$

where C_1 and C_2 are arbitrary constants.

Using the series for $y_1(x)$ and $y_2(x)$, simple calculation gives $y_1(0) = 1$, $y_1'(0) = 0$, $y_2(0) = 0$, and $y_2'(0) = 1$, so the initial conditions $y(0) = 1$, $y'(0) = 0$ will be satisfied if the constants C_1 and C_2 are such that

$$1 = C_1 y_1(0) + C_2 y_2(0) \quad \text{and} \quad 0 = C_1 y_1'(0) + C_2 y_2'(0).$$

This pair of equations has the solution $C_1 = 1$ and $C_2 = 0$, so the solution of the initial value problem becomes

$$y(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^m m!}.$$

Rewriting this as

$$y(x) = \sum_{m=0}^{\infty} \frac{(-x^2/2)^m}{m!},$$

we recognize that the solution is simply $y(x) = \exp(-x^2/2)$, so this series is known to converge for all x.

Finally, to complete our examination of the two linearly independent solutions, let us find the radius of convergence of the second solution $y_2(x)$. The formula for the radius of convergence R based on the ratio test requires all powers of x to be present, whereas the series $y_2(x)$ only contains odd powers of x, so we must modify the series before using the test. All that is necessary is to set $z = x^2$ and to write the series in the form

$$y_2(x) = x \sum_{m=1}^{\infty} \frac{(-1)^m 2^m m!}{(2m+1)!} z^m,$$

for now the radius of convergence of the series in z can be found. The coefficient a_m of z^m is

$$a_m = (-1)^m \frac{2^m m!}{(2m+1)!},$$

so the radius of convergence R is given by

$$R = \lim_{m\to\infty} 1/|a_{m+1}/a_m| = \lim_{m\to\infty} \left| \frac{a_m}{a_{m+1}} \right| = \lim_{m\to\infty} \left(\frac{2^m m!}{(2m+1)!} \frac{(2m+3)!}{2^{m+1}(m+1)!} \right)$$
$$= \lim_{m\to\infty} (2m+3) = \infty.$$

As the series in z has an infinite radius of convergence, so also does the original series involving odd powers of x. This means that the general solution

$$y(x) = C_1 y_1(x) + C_2 y_2(x)$$

is valid for all real x. ■

Legendre's equation

An important application of the power series method of solution is to the **Legendre differential equation**

$$(1-x^2)y'' - 2xy' + \alpha(\alpha+1)y = 0, \tag{10}$$

in which $\alpha \geq 0$ is a real parameter. The equation arises in a variety of applications, but mainly in connection with physical problems in which spherical symmetry is present. It will be seen later that the equation finds its origin in the study of Laplace's equation when expressed in spherical coordinates. Solutions of (10) are called **Legendre functions**, and they are examples of **special functions**, or so-called **higher transcendental functions**, as distinct from elementary functions such as sine, cosine, exponential, and logarithm. We first develop the series solutions for arbitrary $\alpha \geq 0$, and then consider the cases $\alpha = n = 0, 1, 2, \ldots,$ which lead to a special class of polynomial solutions $P_n(x)$ called **Legendre polynomials** in which n is the degree of the polynomial. The important properties of Legendre polynomials will be examined later when the topic of orthogonal functions is introduced.

The coefficients of Legendre's equation are all analytic at the origin and the leading coefficient $(1-x^2)$ only vanishes at $x = \pm 1$, so a power series solution can be expected to exist in the interval $-1 < x < 1$. Substituting (7) to (9) in (10) leads to the equation

$$(1-x^2)\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - 2x\sum_{n=1}^{\infty} na_n x^{n-1} + \alpha(\alpha+1)\sum_{n=0}^{\infty} a_n x^n = 0.$$

Proceeding as in Example 8.4, this can be rewritten as

$$\sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}x^n - \sum_{n=2}^{\infty} n(n-1)a_n x^n - \sum_{n=1}^{\infty} 2na_n x^n + \alpha(\alpha+1)\sum_{n=0}^{\infty} a_n x^n = 0,$$

so equating each coefficient to zero in the usual manner gives the following:
Coefficient of x^0:

$$2a_2 + \alpha(\alpha+1)a_0 = 0,$$

Coefficient of x:

$$6a_3 - 2a_1 + \alpha(\alpha+1)a_1 = 0,$$

Coefficient of x^n for $n \geq 2$:

$$(n+2)(n+1)a_{n+2} - n(n-1)a_n - 2na_n + \alpha(\alpha+1)a_n = 0.$$

Solving the first two equations gives

$$a_2 = -\frac{\alpha(\alpha+1)}{2}a_0 \quad \text{and} \quad a_3 = \frac{[2-\alpha(\alpha+1)]}{6}a_1,$$

whereas the third result gives the recurrence relation

$$a_{n+2} = -\frac{(\alpha-n)(\alpha+n+1)}{(n+2)(n+1)}a_n \quad \text{for } n \geq 2. \tag{11}$$

Straightforward calculations show that the first few coefficients are given by

$$a_2 = -\frac{\alpha(\alpha+1)}{2!}a_0, \quad a_3 = -\frac{(\alpha-1)(\alpha+2)}{3!}a_1,$$
$$a_4 = \frac{(\alpha-2)\alpha(\alpha+1)(\alpha+3)}{4!}a_0, \quad a_5 = \frac{(\alpha-3)(\alpha-1)(\alpha+2)(\alpha+4)}{5!}a_1,$$
$$a_6 = -\frac{(\alpha-4)(\alpha-2)\alpha(\alpha+1)(\alpha+3)(\alpha+5)}{6!}a_0.$$

Thus, the coefficients of the even powers of x are all multiples of a_0, whereas the coefficients of the odd powers of x are all multiples of a_1, where a_0 and a_1 are arbitrary real numbers. Substituting these coefficients into the series

$$y(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots = \sum_{n=0}^{\infty} a_n x^n$$

shows that the general solution of the Legendre differential equation can be written

$$y(x) = a_0 y_1(x) + a_1 y_2(x), \tag{12}$$

where

$$y_1(x) = 1 - \frac{\alpha(\alpha+1)}{2!}x^2 + \frac{(\alpha-2)\alpha(\alpha+1)(\alpha+3)}{4!}x^4 - \cdots, \tag{13}$$

and

$$y_2(x) = x - \frac{(\alpha-1)(\alpha+2)}{3!}x^3 + \frac{(\alpha-3)(\alpha-1)(\alpha+2)(\alpha+4)}{5!}x^5 - \cdots. \tag{14}$$

As the solutions $y_1(x)$ and $y_2(x)$ are not proportional, they must be linearly independent solutions of the Legendre equation (10). We leave as an exercise the task of showing that each series is convergent in the interval $-1 < x < 1$, so the general solution (12) has this same interval of convergence.

Examination of the recurrence relation (11) shows that if $\alpha = n$ is a nonnegative integer, the terms $a_{n+2} = a_{n+4} = a_{n+6} = \cdots$ all vanish. Thus, if $\alpha = n$ is even, the series $y_1(x)$ will reduce to a polynomial of degree n in even powers of x, whereas if $\alpha = n$ is odd the series $y_2(x)$ will reduce to a polynomial of degree n in odd powers of x.

The solution $y(x)$ reduces to the following polynomials when $n = 0, 1, 2, 3, 4$:

Case $n = 0$:

$$y(x) = a_0,$$

Case $n = 1$:

$$y(x) = a_1 x,$$

Case $n = 2$:

$$y(x) = a_0(1 - 3x^2),$$

Case $n = 3$:

$$y(x) = a_1\left(x - \frac{5}{3}x^3\right),$$

Case $n = 4$:

$$y(x) = a_0\left(1 - 10x^2 + \frac{35}{3}x^4\right).$$

When α is a nonnegative integer, after suitable scaling the foregoing polynomials are denoted by $P_n(x)$ and called **Legendre polynomials of degree *n***. The standard scaling adopted involves choosing the arbitrary multiplier of each polynomial such that $P_n(1) = 1$ for $n = 0, 1, 2, \ldots$. When this is done the first few Legendre polynomials become

Legendre polynomials

Even polynomials	*Odd polynomials*
$P_0(x) = 1$	$P_1(x) = x$
$P_2(x) = \frac{1}{2}(3x^2 - 1)$	$P_3(x) = \frac{1}{2}(5x^3 - 3x)$
$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$	$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$

A general expression for $P_n(x)$ can be obtained by writing the recurrence relation (11) in the form

$$a_r = \frac{(r+2)(r+1)}{(r-n)(n+r+1)}a_{r+2} \quad \text{for } r \leq n-2$$

and finding that

$$a_n = \frac{1 \cdot 3 \cdot 4 \cdots (2n-1)}{n!} = \frac{(2n)!}{2^n(n!)^2} \quad \text{for } n = 1, 2, 3, \ldots,$$

in order to make $P_n(1) = 1$. As a result, the following expressions for $P_n(x)$ are obtained.

For *even* polynomials:

$$P_{2n}(x) = \sum_{r=0}^{n}(-1)^r \frac{(4n-2r)!}{2^{2n}r!(2n-r)!(2n-2r)!}x^{2n-2r}, \quad n = 0, 1, 2, \ldots. \tag{15a}$$

For *odd* polynomials:

$$P_{2n+1}(x) = \sum_{r=0}^{n}(-1)^r \frac{(4n-2r+2)!}{2^{2n+1}r!(2n-r+1)!(2n-2r+1)!}x^{2n-2r+1}, \quad n = 0, 1, 2, \ldots. \tag{15b}$$

Two alternative definitions of Legendre polynomials are to be found in Exercises 16 and 18 at the end of this section.

Results (15a, b) provide a general definition for a Legendre polynomial of any order, though when only a few low order polynomials are required it is often more convenient to generate them by means of the following recurrence relation that

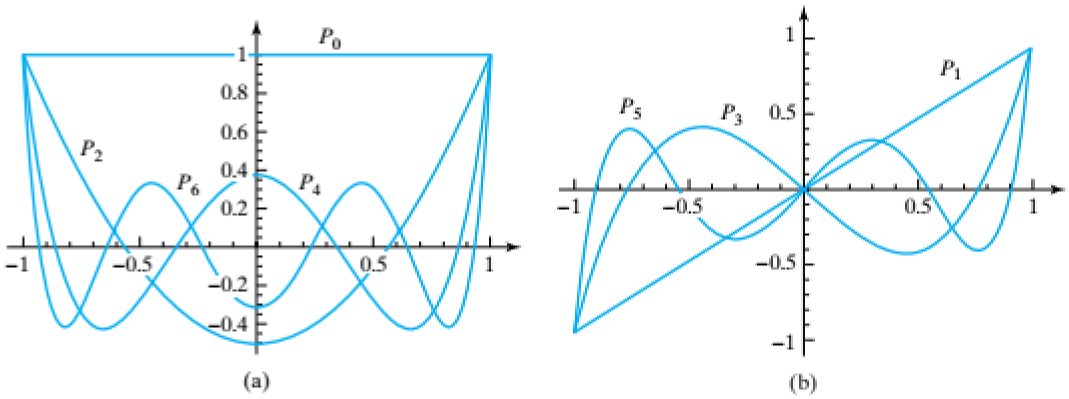


FIGURE 8.1 (a) Even Legendre polynomials. (b) Odd Legendre polynomials.

determines $P_{n+1}(x)$ in terms of $P_n(x)$ and $P_{n-1}(x)$:

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0, \tag{16}$$

recurrence relation for Legendre polynomials

for $n = 1, 2, 3, \dots$. A derivation of this recurrence relation is to be found in Exercise 17 at the end of this section.

As an example of the use of (16) we set $n = 2$ to obtain

$$P_3(x) = \frac{1}{3}[5xP_2(x) - 2P_1(x)],$$

but $P_1(x) = x$ and $P_2(x) = \dfrac{1}{2}(3x^2 - 1)$, so substituting these expressions, we find $P_3(x) = \frac{1}{2}(5x^3 - 3x)$.

Graphs of the first few Legendre polynomials $P_n(x)$ are given in Fig. 8.1.

ADRIEN-MARIE LEGENDRE (1752–1833)
A French mathematician educated at a college in Paris whose remarkable mathematical ability enabled him to be appointed to the position of professor of mathematics at a military school in Paris. His work on the motion of projectiles in a resisting medium won him a prize offered by the Royal Academy in Berlin. He was subsequently appointed professor at the Normal School in Paris and his contributions as an analyst were second only to those of Laplace and Lagrange, who were his contemporaries. In addition to his contributions to the development of the calculus, he made major contributions to the study of elliptic functions.

For more information about Legendre polynomials, and for applications to boundary value problems, see Chapters 5 and 8 of reference [3.7]. Recurrence relations satisfied by Legendre polynomials and other orthogonal polynomials are to be found in Chapter 22 of reference [G.1], and also in Chapter 18 of reference [G.3].

Chebyshev equation

Another important and useful differential equation with a power series solution is the **Chebyshev equation**,

$$(1 - x^2)y'' - xy' + \alpha y = 0. \tag{17}$$

The coefficients are all analytic functions and the leading coefficient $(1 - x^2)$ only vanishes at $x = \pm 1$, so a power series solution can be found in the interval

$-1 \le x \le 1$. Proceeding as with Legendre's equation we find

$$(1-x^2)\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - x\sum_{n=1}^{\infty} na_n x^{n-1} + \alpha\sum_{n=0}^{\infty} a_n x^n = 0,$$

or after a shift of summation index,

$$\sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}x^n - \sum_{n=2}^{\infty} n(n-1)a_n x^n - \sum_{n=1}^{\infty} na_n x^n + \alpha\sum_{n=0}^{\infty} a_n x^n = 0.$$

If we combine summations, this becomes

$$\begin{aligned}&(1\cdot 2a_2 + \alpha a_0) + [2\cdot 3a_3 + (\alpha-1)a_1]x\\ &\quad + \sum_{n=2}^{\infty}\left[(n+1)(n+2)a_{n+2} + (\alpha - n^2)a_n\right]x^n = 0.\end{aligned}$$

Equating the coefficients of each power of x to zero gives

$$a_2 = -\frac{\alpha}{2!}a_0, \quad a_3 = \frac{(1-\alpha)}{3!}a_1,$$

and the recurrence relation

$$a_{n+2} = \frac{(n^2-\alpha)}{(n+1)(n+2)}a_n, \quad n = 2, 3\ldots.$$

Thus,

$$\begin{aligned} a_4 &= \frac{(2^2-\alpha)}{3\cdot 4}a_2 = -\frac{\alpha(2^2-\alpha)}{4!}a_0\\ a_5 &= \frac{(3^2-\alpha)}{4\cdot 5}a_3 = \frac{(1-\alpha)(3^2-\alpha)}{5!}a_1\\ &\ldots\ldots \end{aligned}$$

Using these coefficients in the original power series $y(x) = \sum_{n=0}^{\infty} a_n x^n$ gives the solution of the Chebyshev equation in the form

$$y(x) = a_0 y_0(x) + a_1 y_1(x),$$

where

$$y_0(x) = a_0\left[1 - \frac{\alpha}{2!}x^2 - \frac{\alpha(2^2-\alpha)}{4!}x^4 - \frac{\alpha(2^2-\alpha)(4^2-\alpha)}{6!}x^6 - \cdots\right]$$

and

$$y_1(x) = a_1\left[x + \frac{(1-\alpha)}{3!}x^3 + \frac{(1-\alpha)(3^2-\alpha)}{5!}x^5 + \cdots\right].$$

In applications of this equation to approximation theory, numerical analysis, and elsewhere, it is usual that $\alpha = m^2$, where $m = 0, 1, 2, \ldots$. Inspection of $y_0(x)$ shows that when m is even, the solution reduces to a polynomial of degree m in even powers of x, whereas when m is odd $y_1(x)$ reduces to a polynomial of degree m in odd powers of x.

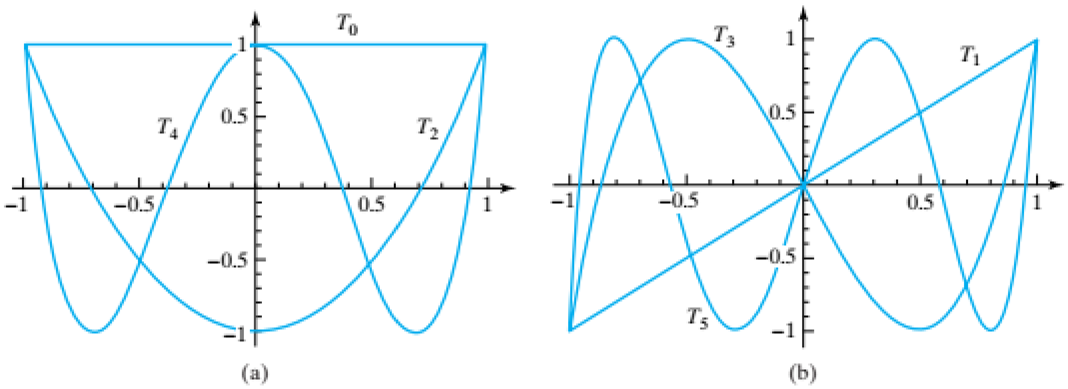


FIGURE 8.2 (a) Even Chebyshev polynomials. (b) Odd Chebyshev polynomials.

As the polynomials are solutions of a homogeneous differential equation, the scale factors for each polynomial can be chosen arbitrarily, so by convention they are chosen such that the term with the largest power of x is positive and the polynomial is free from fractional coefficients. These polynomials are called **Chebyshev polynomials**, and they are denoted by $T_n(x)$. The first six Chebyshev polynomials are:

Chebyshev polynomials

Even polynomials	*Odd polynomials*
$T_0(x) = 1$	$T_1(x) = x$
$T_2(x) = 2x^2 - 1$	$T_3(x) = 4x^3 - 3x$
$T_4(x) = 8x^4 - 8x^2 + 1$	$T_5(x) = 16x^5 - 20x^3 + 5x$

Using the forms for $T_{n+1}(x)$, $T_n(x)$ and $T_{n-1}(x)$ obtained from $y_0(x)$ and $y_1(x)$, it can be shown that Chebyshev polynomials obey the following recurrence relation:

recurrence relation for Chebyshev polynomials

$$T_{n+1}(x) - 2xT_n(x) + T_{n-1}(x) = 0. \tag{18}$$

When used with the polynomials just listed, this recurrence relation is the simplest way of generating higher order polynomials. Graphs of the first six Chebyshev polynomials are shown in Fig. 8.2.

For applications of Chebyshev polynomials to numerical analysis see, for example, references [8.3] to [8.5].

PAFNUTI LIWOWICH CHEBYSHEV (1821–1894)
A distinguished Russian mathematician who was professor of mathematics at the University of Petrograd (now St. Petersburg). He made many contributions to analysis and number theory. There are many variations of the transliteration of his name, the most common probably being Tchebycheff.

Summary

This section showed how to find a series solution, expanded about the origin, of a homogeneous linear second order variable coefficient differential equation with polynomial coefficients, when the solution can be obtained in the form of a general power series with unknown coefficients. By substituting this series into the differential equation, grouping corresponding powers of x, and requiring the coefficient of each power of x to vanish identically, a recurrence relation connecting the unknown coefficients was obtained and used to find the coefficients of the power series in terms of two arbitrary constants a_0 and a_1. The general solution was seen to be the sum of two linearly independent power series with known coefficients, one multiplied by a_0 and the other by a_1. Two important special cases were considered that gave rise to polynomial solutions of the important and useful Legendre and Chebyshev equations.

EXERCISES 8.2

Find the first six terms in the power series expansion of each of the following initial value problems.

1. $y'' + (x - x^2)y' + y = 0$, with $y(0) = 2$, $y'(0) = -3$.
2. $2y'' + xy' + 2(1 + x)y = 0$, with $y(0) = -2$, $y'(0) = 1$.
3. $y'' + (1 + x^2)y' + xy = 0$, with $y(0) = 1$, $y'(0) = -3$.
4. $y'' - 3xy' + 2y = 0$, with $y(0) = 1$, $y'(0) = 1$.
5. $(1 - x^2)y'' + xy' - y = 0$, with $y(0) = 2$, $y'(0) = -1$.
6. $y'' + x^2 y' + 2xy = 0$, with $y(0) = 3$, $y'(0) = -2$.
7. $y'' + 2(1 - x)y' - 3xy = 0$, with $y(0) = 1$, $y'(0) = -1$.
8. $(1 - x)y'' + 2xy' + (1 + x)y = 0$, with $y(0) = 4$, $y'(0) = -2$.
9. $(1 - 2x^2)y'' + 2y' + 3y = 0$, with $y(0) = 1$, $y'(0) = -1$.
10. $(1 + 2x^2)y'' + 3xy' + y = 0$, with $y(0) = 2$, $y'(0) = -2$.
11. $(2x^2 - 1)y'' + (1 + x)y' + 2y = 0$, with $y(0) = 1$, $y'(0) = 4$.
12. $y'' + (1 + 2x)y' + xy = 0$, with $y(2) = 1$, $y'(2) = 0$.
13. $(2 + x)y'' + 3(1 + x)y' + 2y = 0$, with $y(1) = 2$, $y'(1) = -3$.
14. $(x^2 - 2x + 2)y'' + (x - 1)y' - 3y = 0$, with $y(-1) = 1$, $y'(-1) = 2$.
15. $(1 - x)y'' + 2xy' - 2xy = 0$, with $y(2) = 1$, $y'(2) = 5$.
16. An alternative definition of the Legendre polynomial $P_n(x)$ is provided by the formula
$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n,$$
called the **Rodrigues formula**. Use the formula to compute $P_4(x)$ and $P_5(x)$.
17.* Set $u = (x^2 - 1)^n$ and use repeated differentiation of the Rodrigues formula to verify that $P_n(x)$ is a Legendre polynomial by showing it satisfies the Legendre differential equation
$$(1 - x^2)P_n''(x) - 2xP_n'(x) + n(n + 1)P_n(x) = 0.$$
18.* The function
$$G(x, t) = (1 - 2xt + t^2)^{-1/2}$$
is called the **generating function** for Legendre polynomials. It has the property that when expanded as a power series in t the coefficient of t^n is $P_n(x)$, so that
$$G(x, t) = P_0(x) + P_1(x)t + P_2(x)t^2 + \cdots.$$
Set $u = -2xt + t^2$ and expand $(1 + u)^{-1/2}$ by the binomial theorem. Collect all the terms in x multiplying t^5 and hence verify that the coefficient of t^5 is $P_5(x)$.
19.* Show that the generating function defined in Problem 18 satisfies the differential equation
$$(1 - 2xt + t^2)\frac{\partial G}{\partial t} - (x - t)G = 0$$
for arbitrary t. As the result must be an identity in t, the consequence of substituting
$$G(x, t) = P_0(x) + P_1(x)t + P_2(x)t^2 + \cdots$$
into the differential equation must be such that terms in x multiplying each power of t vanish. Collect the terms multiplying t^n, and hence establish the Legendre polynomial recurrence relation
$$(n + 1)P_{n+1}(x) - (2n + 1)xP_n(x) + nP_{n-1}(x) = 0$$
for $n = 1, 2, \ldots$. This result is called the **Bonnet recurrence relation**.
20.* The **electrostatic potential** ϕ at a point in a vacuum distant d from a charge Q is given by $\phi = Q/d$. Use the Legendre polynomial generating function
$$G(r, t) = \frac{1}{(1 - 2rt + t^2)^{1/2}},$$
together with the result from elementary trigonometry
$$r = \left(r_1^2 + r_2^2 - 2r_1 r_2\cos\theta\right)^{1/2},$$
to show that the electrostatic potential at point A due to a charge Q at B in Fig. 8.3 is given by
$$\frac{Q}{r} = \frac{1}{r_2}\left[P_0(\cos\theta) + \left(\frac{r_1}{r_2}\right)P_1(\cos\theta) + \left(\frac{r_1}{r_2}\right)^2 P_2(\cos\theta) + \cdots\right], \text{ for } r_1/r_2 < 1.$$

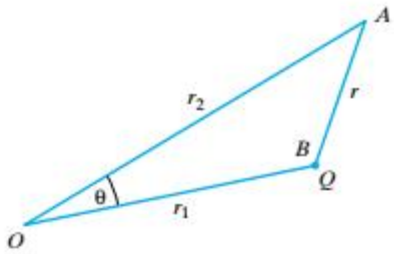


FIGURE 8.3 A point charge Q at B distant r from A.

8.3 Singular Points of Linear Differential Equations

In Section 8.2 the power series method was used to find a solution of a homogeneous variable coefficient differential equation of the form

$$a(x)y'' + b(x)y' + c(x)y = 0. \tag{19}$$

It was seen that the method could be applied about any point x_0 at which the coefficients of the differential equation are analytic and $a(x_0) \neq 0$. Expressed differently, when (19) is written in the standard form

$$y'' + P(x)y' + Q(x)y = 0, \tag{20}$$

with

$$P(x) = \frac{b(x)}{a(x)} \quad \text{and} \quad Q(x) = \frac{c(x)}{a(x)}, \tag{21}$$

the power series method can be applied to develop a solution about any point x_0 at which the functions $P(x)$ and $Q(x)$ are analytic.

regular and singular points

Points where $P(x)$ and $Q(x)$ are analytic are called **regular points** of the differential equation, and points where at least one is not analytic are called **singular points**.

Equation (20) will be said to have a **regular singular point** at x_0 if the functions

$$(x - x_0)P(x) \quad \text{and} \quad (x - x_0)^2 Q(x)$$

are analytic at x_0, and so have Taylor series expansions about x_0. If at least one of these functions is not analytic at x_0, the point will be said to be an **irregular singular point**.

EXAMPLE 8.6

Identify the nature of the singular points of the following equations:

(a) $x^2y'' + xy' + (x^2 - n^2)y = 0$

(b) $(1 - x^2)y'' - 2xy' + n(n+1)y = 0, \quad (n = 0, 1, 2, \ldots)$

(c) $(1-x)y'' + 2(x-1)y' + xy = 0$
(d) $(x-1)^3y'' + 3(x-1)^2y' + y = 0$

Solution

(a) This is *Bessel's equation* of order n in which the functions $P(x) = 1/x$ and $Q(x) = (x^2 - n^2)/x^2$. Neither of these functions is analytic at the origin, so the origin is a singular point of Bessel's equation. However, as the functions $xP(x) = 1$ and $x^2Q(x) = x^2 - n^2$ are both analytic at the origin, it follows that $x = 0$ is a regular singular point of Bessel's equation.

(b) This is *Legendre's equation* of order n in which $P(x) = -2x/(1-x^2)$ and $Q(x) = n(n+1)/(1-x^2)$. Neither of these functions is analytic at $x = \pm 1$, so these points are the singular points of the Legendre equation. Let us consider the singular point at $x = 1$. As the functions

$$(x-1)P(x) = 2x/(1+x) \quad \text{and} \quad (x-1)^2Q(x) = n(n+1)(x-1)/(1+x)$$

are both analytic at $x = 1$, it follows that this is a regular singular point of Legendre's equation. A similar argument shows that $x = -1$ is also a regular singular point of the equation.

(c) In this case $P(x) = -2$ and $Q(x) = x/(1-x)$, and while $P(x)$ is analytic for all x the function $Q(x)$ is not analytic at $x = 1$, so this is a singular point of the equation. The functions $(x-1)P(x) = 2(1-x)$ and $(x-1)^2Q(x) = x(1-x)$ are both analytic at $x = 1$, so $x = 1$ is a regular singular point of the equation.

(d) In this equation $P(x) = 3/(x-1)$ and $Q(x) = 1/(x-1)^3$ and neither function is analytic at $x = 1$, so this is a singular point of the equation. We have

$$(x-1)P(x) = 3 \quad \text{and} \quad (x-1)^2Q(x) = \frac{1}{x-1},$$

and although the first of these functions is analytic for all x, the second is not analytic at $x = 1$, so $x = 1$ is an irregular singular point of the equation. ■

In the next section the power series method will be generalized to arrive at what is called the **Frobenius method**, which always generates two linearly independent solutions about a *regular singular point* of equation (20). As the behavior of solutions in a neighborhood of an irregular singular point can be shown to be very erratic, no further consideration will be given to solutions near such points.

shifting a singular point

Sometimes it is more convenient to consider an equation with a regular singular point located at the origin rather at some other point $x_0 \neq 0$. In such cases a singular point located at x_0 can always be shifted to the origin by making the change of variable $X = x - x_0$, as in Section 8.2.

EXAMPLE 8.7

Shift the singular point of the following equation to the origin:

$$(x-1)^2y'' + 3(x+2)y' + 2y = 0.$$

Solution The equation has a regular singular point at $x = 1$, so we make the variable change $X = x - 1$ and set $y(x) = Y(x-1) = Y(X)$. The equation then becomes

$$X^2Y'' + 3(X+3)Y' + 2Y = 0,$$

with a regular singular point now located at $X = 0$. ■

example showing why no power series solution exists about a singular point

To appreciate why an ordinary power series solution cannot be developed around a regular singular point, it will be sufficient to consider the Cauchy–Euler equation

$$x^2y'' + 3xy' + 2y = 0,$$

which has a regular singular point at the origin. This Cauchy–Euler equation was solved analytically in Example 6.10, where its solution was found to be

$$y(x) = C_1x^{-1}\cos(\ln|x|) + C_2x^{-1}\sin(\ln|x|).$$

The reason that no power series solution exists in this case is seen to be the presence of the factor x^{-1} and the function $\ln|x|$ in the analytical solution, neither of which can be expanded in a power series about the origin.

Summary

The regular and singular points of a general homogeneous second order linear variable coefficient differential equation were defined and illustrated by example. It was shown how, if necessary, a singular point occurring at $x = a$ could be shifted to the origin, and an example was used to demonstrate why an ordinary power series solution cannot be developed around a regular singular point.

EXERCISES 8.3

Identify the nature of the singular points in each of the following equations.

1. $(1-x)^2y'' + 2(x-1)y' + y = 0.$
2. $x^2y'' + 3x^2y' + (1+x^2)y = 0.$
3. $(1+x)^2y'' + 2y' + y = 0.$
4. $xy'' + (1-x)y' + ny = 0 \quad (n > 0).$
5. $(x+4)^3y'' + 2(x+4)y' + xy = 0.$
6. $(x^2-4)y'' + (x+3)y' - 5(x+1)y = 0.$
7. $(3-x)^2y'' + 4y' + \cos x(3-x^2)y = 0.$
8. $x^2y'' + 8y' + 3xy = 0.$

8.4 The Frobenius Method

A generalization of the power series method that was introduced by Frobenius (1849–1917) enables a solution of a homogeneous linear differential equation to be developed about a regular singular point. He considered the differential equation

$$a(x)y'' + b(x)y' + c(x)y = 0, \tag{22}$$

and established the following result that is stated without proof.

THEOREM 8.1

Frobenius theorem Let x_0 be a regular singular point of (22). Then, in some interval $0 < x - x_0 < d$, the equation will always possess at least one solution of the form

$$y(x) = (x-x_0)^c\left(a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots\right)$$

$$= (x-x_0)^c\sum_{n=0}^{\infty} a_n(x-x_0)^n,$$

the Frobenius theorem and method of solution

where $a_0 \neq 0$ and c is a real or complex number. A second linearly independent solution of similar form will exist that may contain a logarithmic term, though with a different value of c and some other coefficients $b_0, b_1, b_2, \ldots$ in place of the coefficients $a_0, a_1, a_2, \ldots$. Taken together, these two solutions form a basis of solutions for the differential equation. ■

GEORG FERDINAND FROBENIUS (1849–1917)
A German mathematician whose main research was in group theory and analysis. He worked in Zurich and Berlin and published his method for the series solution of linear ordinary differential equations in 1873.

For simplicity, and because of their frequent occurrence, in what follows we will develop the Frobenius method in terms of a slightly less general class of equations by setting $a(x) = x^2$ in (22). So we will consider the equation

$$x^2 y'' + b(x)y' + c(x)y = 0, \tag{23}$$

and write it in the standard form

$$y'' + P(x)y' + Q(x)y = 0, \tag{24}$$

where

$$P(x) = \frac{p(x)}{x} \quad \text{and} \quad Q(x) = \frac{q(x)}{x^2}, \tag{25}$$

and assume that $p(x)$ and $q(x)$ are analytic functions at $x = 0$. So we will only consider equations of the form (24) with regular singular points at the *origin*.

To determine the exponent c in Theorem 8.1 we substitute a solution of the form

$$y(x) = x^c \sum_{n=0}^{\infty} a_n x^n \tag{26}$$

into equation (24), where c is to be determined along with the coefficients a_n. When making this substitution we will need to use the following results obtained by differentiation of (26):

$$y'(x) = ca_0 x^{c-1} + (c+1)a_1 x^c + (c+2)a_2 x^{c+1} + \cdots = \sum_{n=0}^{\infty} (n+c)a_n x^{n+c-1} \tag{27}$$

and

$$\begin{aligned} y''(x) &= c(c-1)a_0 x^{c-2} + (c+1)ca_1 x^{c-1} + (c+2)(c+1)a_2 x^c + \cdots \\ &= \sum_{n=0}^{\infty} (n+c)(n+c-1)a_n x^{n+c-2}. \end{aligned} \tag{28}$$

As the functions $p(x)$ and $q(x)$ are assumed to be analytic at the origin, they can be expanded as the Maclaurin series

$$p(x) = p_0 + p_1 x + p_2 x^2 + \cdots \quad \text{and} \quad q(x) = q_0 + q_1 x + q_2 x^2 + \cdots. \tag{29}$$

Substituting (27) to (29) into (24) leads to the result

$$\begin{aligned} &x^{c-2}[c(c-1)a_0 + (c+1)ca_1 x + \cdots] \\ &\quad + \left(p_0 + p_1 x + p_2 x^2 + \cdots\right)x^{c-2}(ca_0 + (c+1)a_1 x + \cdots) \\ &\quad + x^{c-2}\left(q_0 + q_1 x + q_2 x^2 + \cdots\right)\left(a_0 + a_1 x + a_2 x^2 + \cdots\right) = 0. \end{aligned}$$

If (26) is to be a solution of (24), the coefficient of each power of x in this last result must vanish to make it an identity. Collecting terms involving the same power of x and equating their coefficients to zero will lead to a sequence of equations connecting the coefficients a_n in (26), and equating the coefficient of the lowest power of x to zero will give an equation from which c can be determined.

The lowest power of x in the preceding result is x^{c-2}, so collecting terms involving x^{c-2} and equating the coefficient of x^{c-2} to zero gives

$$[c(c-1) + p_0 c + q_0]a_0 = 0.$$

As Theorem 8.1 requires $a_0 \neq 0$, it follows that c is determined by the equation

$$c(c-1) + p_0 c + q_0 = 0. \tag{30}$$

indicial equation

This equation is called the **indicial equation** associated with differential equation (24), because it determines the permissible values of the index c to be used in the solution given in Theorem 8.1.

The indicial equation of differential equation (24) can be constructed without the need to make the substitution (26), because it is easily seen that

$$p_0 = \lim_{x\to 0}[xP(x)] \quad \text{and} \quad q_0 = \lim_{x\to 0}[x^2Q(x)]. \tag{31}$$

For the class of equations of type (24) that all have a regular singular point at the origin, the appropriate form of the Frobenius theorem follows from Theorem 8.1 if we set $x_0 = 0$.

It is important to notice that for a general equation (22) in which $a(x) \neq x^2$ the indicial equation does *not* take the form given in (30). When this situation arises the indicial equation must be obtained by substituting (26) into (22) and equating to zero the coefficient of the lowest power of x that occurs in the expansion.

As the indicial equation is a quadratic equation in c, the following relationships between its roots c_1 and c_2 are possible:

(a) Roots c_1 and c_2 are real and distinct and do not differ by an integer
(b) Roots c_1 and c_2 are real and differ by an integer
(c) Roots c_1 and c_2 are real and equal
(d) Roots c_1 and c_2 are complex conjugates

The reason for identifying these different cases is to be found in the following theorem, which is stated without proof in terms of a differential equation with a regular singular point located at the origin (see references [3.3] and [3.5]).

THEOREM 8.2

Forms of Frobenius solution depending on the nature of c_1 and c_2 Let a differential equation of the form

$$x^2y'' + x[xP(x)]y' + [x^2Q(x)]y = 0$$

have a regular singular point at $x = 0$. Let $xP(x)$ and $x^2Q(x)$ each be capable of expansion as convergent power series in an interval $|x| < d$, where $d > 0$ is the smaller of the two radii of convergence, and suppose that

$$p_0 = \lim_{x \to 0}[xP(x)] \quad \text{and} \quad q_0 = \lim_{x \to 0}[x^2 Q(x)].$$

Then in terms of the exponent c in (26), and the coefficients p_0 and q_0, the indicial equation for the differential equation is

$$c(c-1) + p_0 c + q_0 = 0,$$

with two roots c_1 and c_2 that may be real or complex conjugates.

The two linearly independent solutions of the differential equation that exist depend on the relationship between the roots of the indicial equation, and they take the following forms.

different forms of Frobenius solution and examples

Case (a) Real roots with $c_1 > c_2$ and $c_1 - c_2$ neither zero nor a positive integer

In the intervals $-d < x < 0$ and $0 < x < d$ the differential equation has two linearly independent solutions of the form

$$y_1(x) = |x|^{c_1}\left[1 + \sum_{n=1}^{\infty} a_n x^n\right] \quad \text{and} \quad y_2(x) = |x|^{c_2}\left[1 + \sum_{n=1}^{\infty} b_n x^n\right],$$

where the coefficients a_n are obtained by substituting $c = c_1$ in the recurrence relation connecting coefficients and then setting $a_0 = 1$, and the coefficients b_n are obtained in similar fashion by substituting $c = c_2$ in the recurrence relation, replacing a_n by b_n and setting $b_0 = 1$.

Case (b) Real roots with $c_1 - c_2$ equal to a positive integer

In the intervals $-d < x < 0$ and $0 < x < d$ the differential equation has two linearly independent solutions of the form

$$y_1(x) = |x|^{c_1}\left[1 + \sum_{n=1}^{\infty} a_n x^n\right] \quad \text{and} \quad y_2(x) = A y_1(x) \ln|x| + |x|^{c_2} \sum_{n=1}^{\infty} \beta_n x^n,$$

where the coefficients a_n are determined as in Case (a), and the coefficients A and β_n are found by substituting $y(x) = y_2(x)$ in the differential equation. Some differential equations for which $c_1 - c_2$ is a positive integer have no logarithmic term in their solution $y_2(x)$, in which case $A = 0$.

Case (c) Real roots with $c_1 = c_2$

In the intervals $-d < x < 0$ and $0 < x < d$ the differential equation has two linearly independent solutions of the form

$$y_1(x) = |x|^{c_1}\left[1 + \sum_{n=1}^{\infty} a_n x^n\right] \quad \text{and} \quad y_2(x) = y_1(x) \ln|x| + |x|^{c_1} \sum_{n=1}^{\infty} \alpha_n x^n,$$

where the coefficients a_n are determined as in Case (a), and the coefficients α_n are found by substituting $y(x) = y_2(x)$ into the differential equation.

Case (d) Complex conjugate roots

If $c_1 = \lambda + i\mu$ and $c_2 = \lambda - i\mu$ with $\mu \neq 0$, then in the intervals $-d < x < 0$ and $0 < x < d$ the two linearly independent solutions of the differential equation are the real and imaginary parts of

$$y(x) = |x|^{\lambda + i\mu}\left[1 + \sum_{n=1}^{\infty} a_n x^n\right],$$

where the coefficients a_n are determined as in Case (a). ■

It is important to recognize that the solutions in cases (a) to (d) of Theorem 8.2 all lie in intervals of the form $0 < x < d$ that do *not* contain the origin. A solution in the interval $-d < x < 0$ can be obtained from the above results by replacing x by $-x$ and, depending on the relationship between the roots c_1 and c_2, seeking a solution in the manner indicated in the illustrative examples that follow.

Case (a) Roots c_1 and c_2 Are Distinct and Do Not Differ by an Integer

EXAMPLE 8.8 Find the solution of

$$2xy'' + (x+1)y' + y = 0$$

in some interval $0 < x < d$.

Solution As the coefficient of y'' vanishes at $x = 0$ the origin must be a singular point of this equation. When the differential equation is written in standard form we find that $P(x) = (x+1)/(2x)$ and $Q(x) = 1/(2x)$, so $p_0 = \lim_{x\to 0} xP(x) = 1/2$ and $q_0 = \lim_{x\to 0} x^2 Q(x) = 0$, showing that the origin is a regular singular point of the differential equation.

From (30) the indicial equation is seen to be

$$c(c-1) + \frac{1}{2}c = 0, \quad \text{or} \quad c\left(c - \frac{1}{2}\right) = 0,$$

showing that the permissible values of c are $c = 0$ and $c = 1/2$. As these values of c are distinct and do not differ by an integer, the solution will be of the type given in Theorem 8.2(a).

Setting

$$y(x) = \sum_{n=0}^{\infty} a_n x^{n+c}$$

and substituting into the differential equation in the usual way leads to the result

$$2\sum_{n=0}^{\infty}(n+c)(n+c-1)a_n x^{n+c-1} + \sum_{n=0}^{\infty}(n+c)a_n x^{n+c} + \sum_{n=0}^{\infty}(n+c)a_n x^{n+c-1}$$
$$+ \sum_{n=0}^{\infty} a_n x^{n+c} = 0.$$

Shifting the summation index in the first and third summations gives

$$2\sum_{n=-1}^{\infty}(n+c+1)(n+c)a_{n+1}x^{n+c}+\sum_{n=0}^{\infty}(n+c)a_nx^{n+c}+\sum_{n=-1}^{\infty}(n+c+1)a_{n+1}x^{n+c}$$
$$+\sum_{n=0}^{\infty}a_nx^{n+c}=0,$$

and, finally, combining terms we arrive at the result

$$\sum_{n=-1}^{\infty}[2(n+c+1)(n+c)+(n+c+1)]a_{n+1}x^{n+c}+\sum_{n=0}^{\infty}(n+c+1)a_nx^{n+c}=0.$$

Separating out the term corresponding to $n=-1$ allows this to be written

$$[2c(c-1)+c]a_0x^{c-1}+\sum_{n=0}^{\infty}\{[2(n+c+1)(n+c)+(n+c+1)]a_{n+1}$$
$$+(n+c+1)a_n\}x^{n+c}=0.$$

To proceed further we must now equate to zero the coefficient of each power of x. Equating to zero the coefficient of x^{c-1} simply gives the indicial equation, but equating to zero the coefficient of x^{n+c} for $n=0,1,2,\ldots$ gives

$$(n+c+1)(2n+2c+1)a_{n+1}+(n+c+1)a_n=0.$$

As $n+c+1\neq 0$ this recurrence relation can be written

$$a_{n+1}=-\frac{a_n}{2n+2c+1}.$$

Starting with the value $c=0$, we find that

$$a_{n+1}=-\frac{a_n}{2n+1},$$

so

$$a_1=-a_0,\quad a_2=-\frac{a_1}{3}=\frac{a_0}{3},\quad a_3=-\frac{a_2}{5}=-\frac{a_0}{3\cdot 5},\quad a_4=-\frac{a_3}{7}=\frac{a_0}{3\cdot 5\cdot 7},$$
$$a_5=-\frac{a_4}{9}=-\frac{a_0}{3\cdot 5\cdot 7\cdot 9},\quad a_6=-\frac{a_5}{11}=\frac{a_0}{3\cdot 5\cdot 7\cdot 9\cdot 11},\ldots.$$

Examination of a_5 and a_6 shows they can be written

$$a_5=-\frac{2\cdot 4\cdot 6\cdot 8}{9!}a_0=-\frac{2^4\cdot 4!}{(2\cdot 4+1)!}a_0$$

and

$$a_6=\frac{2^5\cdot 5!}{(2\cdot 5+1)!}a_0.$$

These expressions suggest that the coefficient of the general term in the series is

$$a_{n+1}=\frac{(-1)^{n+1}2^n n!}{(2n+1)!}a_0\quad \text{for } n=0,1,2,\ldots,$$

and this is easily verified by mathematical induction. As we are considering the case in which $c = 0$, it follows from Theorem 8.2(a) that for some $d_1 > 0$ one solution is

$$y(x) = a_0\left[1 + \sum_{n=0}^{\infty}\frac{(-1)^{n+1}2^n n!}{(2n+1)!}x^{n+1}\right].$$

As the constant $a_0 \neq 0$ is arbitrary, we set $a_0 = 1$ and take for a fundamental solution of the differential equation

$$y_1(x) = 1 + \sum_{n=0}^{\infty}\frac{(-1)^{n+1}2^n n!}{(2n+1)!}x^{n+1} \quad \text{for } 0 < x < d_1.$$

A second fundamental (linearly independent) solution follows by using the other value $c = 1/2$, for which the recurrence relation becomes

$$a_{n+1} = -\frac{a_n}{2n+2}.$$

Using this result and recognizing that the coefficients a_n are not the same as the ones in $y_1(x)$, we find that

$$a_1 = -\frac{a_0}{2}, \quad a_2 = -\frac{a_1}{2\cdot 2} = \frac{a_0}{2^2\cdot 2!}, \quad a_3 = -\frac{a_2}{2\cdot 3} = -\frac{a_0}{2^3\cdot 3!},$$
$$a_4 = -\frac{a_3}{2\cdot 4} = \frac{a_0}{2^4\cdot 4!}, \ldots .$$

This pattern of coefficients suggests that the coefficient of the general term in the series is

$$a_n = \frac{(-1)^n}{2^n n!}a_0,$$

and this also is easily verified by using an inductive argument. Setting the arbitrary constant $a_0 = 1$, it follows from Theorem 8.2(a) that for some $d_2 > 0$ a second fundamental solution is given by

$$y_2(x) = x^{1/2}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n n!}x^n = x^{1/2}e^{-x/2}, \quad \text{for } 0 < x < d_2.$$

The solutions $y_1(x)$ and $y_2(x)$ form a basis for solutions of the differential equation in an interval of the form $0 < x < d$, where $d = \min\{d_1, d_2\}$. Thus, the general solution is

$$y(x) = C_1 y_1(x) + C_2 y_2(x), \quad \text{for } 0 < x < d,$$

where C_1 and C_2 are arbitrary constants. The value of d is

$$d = \min\{R_1, R_2\},$$

where R_1 and R_2 are the radii of convergence of the series solutions for $y_1(x)$ and $y_2(x)$, respectively. In this case $R_1 = R_2 = \infty$, so the general solution is valid for $x > 0$. ■

Case (b) Roots c_1 and c_2 Are Real and Differ by an Integer

EXAMPLE 8.9 Find the solution of

$$x^2y'' + x(2+x)y' - 2y = 0$$

in some interval $0 < x < d$.

Solution The equation has a singular point at the origin, and writing it in standard form shows that $P(x) = (2 + x)/x$ and $Q(x) = -2/x^2$. Thus,

$$p_0 = \lim_{x\to 0} xP(x) = 2 \quad \text{and} \quad q_0 = \lim_{x\to 0} x^2 Q(x) = -2,$$

so the equation has a regular singular point at the origin. It follows from (30) that the indicial equation is

$$c(c-1) + 2c - 2 = 0, \quad \text{or} \quad c^2 + c - 2 = 0.$$

The permissible values of c are thus $c = -2$ and $c = 1$, and these differ by an integer.

Substituting the series

$$y(x) = \sum_{n=0}^{\infty} a_n x^{n+c}$$

into the differential equation gives

$$\sum_{n=0}^{\infty}(n+c)(n+c-1)a_n x^{n+c} + 2\sum_{n=0}^{\infty}(n+c)a_n x^{n+c} + \sum_{n=0}^{\infty}(n+c)a_n x^{n+c+1}$$
$$-2\sum_{n=0}^{\infty} a_n x^{n+c} = 0.$$

Shifting the index in the third summation so it starts from $n = 1$ and separating out the terms multiplied by x^c enables the equation to be written

$$a_0(c^2 + c - 2)x^c + \sum_{n=1}^{\infty}\{[(n+c)(n+c+1) - 2]a_n + (n+c-1)a_{n-1}\}x^{n+c} = 0.$$

Proceeding as usual and equating the coefficient of x^c to zero simply gives the indicial equation, whereas equating the coefficient of x^{n+c} to zero gives the recurrence relation

$$a_n = \frac{(n+c-1)}{[2-(n+c)(n+c+1)]}a_{n-1} \quad \text{for } n = 1, 2, \ldots.$$

Considering the larger root $c = 1$, as required by Theorem 8.2(b), we find that

$$a_n = \frac{n}{[2-(1+n)(2+n)]}a_{n-1} \quad \text{for } n = 1, 2, \ldots.$$

So the first few coefficients are

$$a_1 = -\frac{a_0}{4}, \quad a_2 = \frac{2}{[2-3\cdot 4]}a_1 = \frac{a_0}{4\cdot 5}, \quad a_3 = -\frac{3a_2}{[2-4\cdot 5]} = -\frac{a_0}{4\cdot 5\cdot 6},$$
$$a_4 = \frac{4a_3}{[2-5\cdot 6]} = \frac{a_0}{4\cdot 5\cdot 6\cdot 7}, \ldots.$$

As $c = 1$, setting the arbitrary constant $a_0 = 1$, it follows from Theorem 8.2(b) that for some $d_1 > 0$ a fundamental solution of the differential equation is

$$y_1(x) = x\left(1 - \frac{x}{4} + \frac{x^2}{4\cdot 5} - \frac{x^3}{4\cdot 5\cdot 6} + \frac{x^4}{4\cdot 5\cdot 6\cdot 7} - \cdots\right),$$

or

$$y_1(x) = x - \frac{x^2}{4} + \frac{x^3}{4 \cdot 5} - \frac{x^4}{4 \cdot 5 \cdot 6} + \frac{x^5}{4 \cdot 5 \cdot 6 \cdot 7} - \cdots,$$

with $0 < x < d$.

Theorem 8.2(b) asserts that, corresponding to the smaller root $c = -2$, a second fundamental solution is of the form

$$\begin{aligned} y_2(x) &= Cy_1(x)\ln x + x^{-2}\sum_{n=0}^{\infty} b_n x^n \\ &= Cy_1(x)\ln x + \sum_{n=0}^{\infty} b_n x^{n-2}. \end{aligned}$$

To determine C and the coefficients b_n, we substitute this solution into the original differential equation, and because the result must be an identity in x, the coefficient of each power of x must vanish.

Differentiation of the foregoing result gives

$$y_2' = Cy_1'(x)\ln x + \frac{Cy_1(x)}{x} + \sum_{n=0}^{\infty}(n-2)b_n x^{n-3}$$

and

$$y_2''(x) = Cy_1''(x)\ln x + \frac{2Cy_1'(x)}{x} - \frac{Cy_1(x)}{x^2} + \sum_{n=0}^{\infty}(n-2)(n-3)b_n x^{n-4}.$$

Substituting these results into the differential equation and collecting terms leads to the result

$$\begin{aligned} &[x^2 y_1''(x) + x(2+x)y_1'(x) - 2y_1(x)]C\ln x + C[y_1(x) + xy_1(x) + 2xy_1'(x)] \\ &+ \sum_{n=0}^{\infty}(n-3)(n-2)b_n x^{n-2} + \sum_{n=0}^{\infty}2(n-2)b_n x^{n-2} + \sum_{n=0}^{\infty}(n-2)b_n x^{n-1} \\ &- \sum_{n=0}^{\infty}2b_n x^{n-2} = 0. \end{aligned}$$

The coefficient of the logarithmic term vanishes, because $y_1(x)$ is a solution of the differential equation, so the equation simplifies to

$$\begin{aligned} &C[y_1(x) + xy_1(x) + 2xy_1'(x)] \\ &\quad + \sum_{n=0}^{\infty}(n-3)(n-2)b_n x^{n-2} + \sum_{n=0}^{\infty}2(n-2)b_n x^{n-2} + \sum_{n=0}^{\infty}(n-2)b_n x^{n-1} \\ &\quad - \sum_{n=0}^{\infty}2b_n x^{n-2} = 0. \end{aligned}$$

The terms corresponding to $n = 0$ cancel, and after shifting the summation index in the third summation, we have

$$C[y_1(x) + xy_1(x) + 2xy_1'(x)] + \sum_{n=1}^{\infty}(n-3)(nb_n + b_{n-1})x^{n-2} = 0.$$

To find the form of the first group of terms $C[y_1(x) + xy_1(x) + 2xy_1'(x)]$, we must use the series solution for $y_1(x)$. As

$$y_1(x) = x - \frac{x^2}{4} + \frac{x^3}{4 \cdot 5} - \frac{x^4}{4 \cdot 5 \cdot 6} + \cdots,$$

differentiation gives

$$y_1'(x) = 1 - \frac{x}{2} + \frac{3x^2}{20} - \frac{x^3}{30} + \cdots,$$

and so

$$C[y_1(x) + xy_1(x) + 2xy_1'(x)] = 3Cx - \frac{Cx^2}{4} + \frac{Cx^3}{10} - \frac{Cx^4}{40} + \cdots.$$

Using this result in the equation and expanding the first few terms in the summation involving the unknown coefficients b_n shows that

$$\left(3Cx - \frac{Cx^2}{4} + \frac{Cx^3}{10} - \frac{Cx^4}{40} + \cdots\right) - (2b_1 + 2b_0)\frac{1}{x} - (2b_2 + b_1) + (4b_4 + b_3)x^2$$
$$+ (10b_5 + 2b_4)x^3 + (18b_6 + 3b_5)x^4 + (28b_7 + 4b_6)x^5 + (40b_8 + 5b_7)x^6 + \cdots = 0.$$

If we now equate to zero the coefficient of each power of x, we find that

$$b_1 = -b_0, \quad b_2 = -\frac{1}{2}b_1 = \frac{1}{2}b_0, \quad C = 0, \quad b_4 = -\frac{1}{4}b_3,$$
$$b_5 = -\frac{1}{5}b_4 = \frac{1}{4 \cdot 5}b_3, \quad b_6 = -\frac{1}{6}b_5 = -\frac{1}{4 \cdot 5 \cdot 6}b_3, \ldots.$$

The condition $C = 0$ shows that in this case the second linearly independent solution $y_2(x)$ does *not* contain a logarithmic term. The terms b_1 and b_2 are determined as multiples of b_0, and from Theorem 8.2(b) $b_0 \neq 0$, whereas for $n > 3$ all of the terms b_n are seen to be multiples of b_3, which is arbitrary because no equation connects it with b_0. Thus, the solution that has been generated appears to contain *two* arbitrary constants instead of the *one* that would have been expected. Substituting the b_n into the general form of the solution, which with $C = 0$ has reduced to

$$y_2(x) = \sum_{n=0}^{\infty} b_n x^{n-2},$$

gives

$$y_2(x) = b_0\left(\frac{1}{x^2} - \frac{1}{x} + \frac{1}{2}\right) + b_3x\left(1 - \frac{x}{4} + \frac{x^2}{4 \cdot 5} - \frac{x^3}{4 \cdot 5 \cdot 6} + \frac{x^4}{4 \cdot 5 \cdot 6 \cdot 7} - \cdots\right).$$

The apparent incompatibility caused by the introduction of the two arbitrary constants b_0 and b_3 is now resolved, because the series multiplied by b_3 is simply the first linearly independent solution $y_1(x)$. So, in this case, when seeking the second linearly independent solution we have, in fact, generated a linear combination of the first linearly independent solution $y_1(x)$ and another linearly independent solution given by the expression

$$\frac{1}{x^2} - \frac{1}{x} + \frac{1}{2}.$$

Accordingly, we set $b_3 = 0$ and $b_0 = 1$, and take for the second linearly independent solution

$$y_2(x) = \frac{1}{x^2} - \frac{1}{x} + \frac{1}{2},$$

and since only three terms are involved we see that $y_2(x)$ is defined for $x > 0$.

When closed form solutions such as $y_2(x)$ are obtained, they should always be checked by substitution into the differential equation, and in this case it is easy to check that $y_2(x)$ is, indeed, a solution.

It is a simple matter to show the radius of convergence of the series solution $y_1(x)$ is infinite, so solutions $y_1(x)$ and $y_2(x)$ form a basis for the solution of the differential equation whose general solution is

$$y(x) = C_1 y_1(x) + C_2 y_2(x), \quad \text{for } x > 0,$$

where C_1 and C_2 are arbitrary constants. ■

Case (c) Equal Real Roots $c_1 = c_2$

EXAMPLE 8.10 Find the solution of

$$x^2 y'' + (x^2 - x)y' + y = 0,$$

in some interval $0 < x < d$.

Solution This equation has a singular point at the origin, and when expressed in standard form we see that $P(x) = (x-1)/x$ and $Q(x) = 1/x^2$, so

$$p_0 = \lim_{x\to 0} xP(x) = -1 \quad \text{and} \quad q_0 = \lim_{x\to 0} x^2 Q(x) = 1.$$

Thus, the origin is a regular singular point, and from (30) the indicial equation is seen to be

$$c(c-1) - c + 1 = 0, \quad \text{or} \quad (c-1)^2 = 0,$$

so the roots are $c = 1$ (twice).

Substituting the series

$$y(x) = \sum_{n=0}^{\infty} a_n x^{n+c}$$

into the differential equation gives

$$\sum_{n=0}^{\infty}(n+c)(n+c-1)a_n x^{n+c} + \sum_{n=0}^{\infty}(n+c)a_n x^{n+c+1} - \sum_{n=0}^{\infty}(n+c)a_n x^{n+c} + \sum_{n=0}^{\infty} a_n x^{n+c} = 0.$$

Shifting the summation index in the second summation allows it to be written

$$\sum_{n=1}^{\infty}(n+c-1)a_{n-1}x^{n+c},$$

so using this in the preceding equation and separating out the terms corresponding to $n = 0$ we find that

$$a_0[c(c-1) - c + 1]x^c + \sum_{n=1}^{\infty}\{[(n+c)(n+c-2)+1]a_n + (n+c-1)a_{n-1}\}x^{n+c} = 0.$$

As usual, equating the coefficient of x^c to zero gives the indicial equation, and equating the coefficient of x^{n+c} to zero gives the recurrence relation

$$[(n+c)(n+c-2)+1]a_n = -(n+c-1)a_{n-1} \quad \text{for } n = 1, 2, \ldots.$$

Setting $c = 1$ this becomes

$$a_n = -a_{n-1}/n,$$

so

$$a_1 = -a_0, \quad a_2 = -\frac{1}{2}a_0 = \frac{1}{2!}a_0, \quad a_3 = -\frac{1}{3}a_2 = \frac{1}{3!}a_0$$

and, in general,

$$a_n = \frac{(-1)^n}{n!} \quad \text{for } n = 0, 1, 2, \ldots.$$

Setting the arbitrary constant $a_0 = 1$ gives as a fundamental solution of the equation

$$y_1(x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}x^{n+1} = xe^{-x}.$$

The series for e^{-x} converges for $x > 0$, so this result is valid for all $x > 0$.

Continuing, we now illustrate two different methods by which a second linearly independent solution may be found.

Method 1. As the form of solution $y_1(x)$ is particularly simple, we will make use of result (35) of Section 6.3 that asserts that if $y_1(x)$ is a solution of the equation

$$y'' + P(x)y' + Q(x)y = 0,$$

an example using the reduction of order method

then a second linearly independent solution is given by the **reduction of order** formula

$$y_2(x) = y_1(x)\int \frac{\exp[-\int P(x)dx]}{[y_1(x)]^2}dx.$$

Substituting for $y_1(x)$ and $P(x)$ gives

$$\int P(x)dx = \int \frac{(x-1)}{x}dx = x - \ln x, \quad \text{so } \exp\left[-\int P(x)dx\right] = xe^{-x}.$$

Thus,

$$y_2(x) = y_1(x)\int \frac{xe^{-x}}{x^2e^{-2x}}dx = y_1(x)\int \frac{e^x}{x}dx.$$

To integrate this result we replace e^x by its series expansion and integrate term by term to obtain

$$y_2(x) = xe^{-x} \int \left(\frac{1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots}{x} \right) dx$$

$$= xe^{-x} \left(\ln x + x + \frac{x^2}{4} + \frac{x^3}{18} + \frac{x^4}{96} + \frac{x^5}{600} + \cdots \right).$$

In order to compare this method with the one that is to follow, we rewrite this result by replacing e^{-x} by the first few terms of its series expansion to give

$$y_2(x) = xe^{-x} \ln x + x \left(1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots \right) \left(x + \frac{x^2}{4} + \frac{x^3}{18} + \frac{x^4}{96} + \frac{x^5}{600} + \cdots \right).$$

Multiplying the two series together then shows that for some d_2

$$y_2(x) = xe^{-x} \ln x + \left(x^2 - \frac{3x^3}{4} + \frac{11x^4}{36} - \frac{25x^5}{288} + \cdots \right), \quad \text{for } 0 < x < d_2,$$

where d_2 is the radius of convergence of the bracketed series.

Method 2. Theorem 8.2(c) asserts that the second linearly independent solution has the form

$$y_2(x) = y_1(x) \ln\, x + x^2 \sum_{n=0}^{\infty} b_n x^n = y_1(x) \ln x + \sum_{n=0}^{\infty} b_n x^{n+2}.$$

Substituting this result into the differential equation and collecting terms gives

$$\begin{aligned}
&[x^2 y_1''(x) + (x^2 - x) y_1'(x) + y_1(x)] \ln\, x + 2x y_1'(x) + x y_1(x) - 2y_1(x) \\
&+ \sum_{n=0}^{\infty} (n+2)(n+1) b_n x^{n+2} + \sum_{n=0}^{\infty} (n+2) b_n x^{n+3} - \sum_{n=0}^{\infty} (n+2) b_n x^{n+2} \\
&+ \sum_{n=0}^{\infty} b_n x^{n+2} = 0.
\end{aligned}$$

Notice that the logarithmic term has vanished because $y_1(x)$ is a solution of the differential equation.

Shifting the summation index in the second summation, we obtain

$$\begin{aligned}
&2x y_1'(x) + x y_1(x) - 2y_1(x) + \sum_{n=0}^{\infty} (n+2)(n+1) b_n x^{n+2} \\
&+ \sum_{n=1}^{\infty} (n+1) b_{n-1} x^{n+2} - \sum_{n=1}^{\infty} (n+2) b_n x^{n+2} + \sum_{n=0}^{\infty} b_n x^{n+2} = 0.
\end{aligned}$$

Separating out the terms corresponding to $n = 0$ allows this to be written as

$$2x y_1'(x) + x y_1(x) - 2y_1(x) + b_0 x^2 + \sum_{n=1}^{\infty} (n+1)[(n+1) b_n + b_{n-1}] x^{n+2} = 0.$$

The terms involving $y_1(x)$ are now obtained by differentiation of the series

$$y_1(x) = xe^{-x} = x - x^2 + x^3/3 - x^4/6 + x^5/24 - \cdots,$$

leading to

$$2xy_1'(x) + xy_1(x) - 2y_1(x) = -x^2 + x^3 - x^4/2 + x^5/6 - x^6/24 + \cdots.$$

Using this result in the above equation and expanding the terms involving b^n gives

$$(-x^2 + x^3 - x^4/2 + x^5/6 - x^6/24 + \cdots) + b_0x^2 + 2(2b_1 + b_0)x^3$$
$$+ 3(3b_2 + b_1)x^4 + 4(4b_3 + b_2)x^5 + 5(5b_4 + b_3)x^6 + \cdots = 0.$$

Finally, equating the coefficients of powers of x to zero gives

$$b_0 - 1 = 0, \quad 4b_1 + b_0 + 1 = 0, \quad 9b_2 + 3b_1 - 1/2 = 0, \ldots.$$

so that

$$b_0 = 1, \quad b_1 = -3/4, \quad b_2 = 11/36, \quad b_3 = -25/288, \ldots.$$

Substituting these coefficients into the general form of the solution again produces the second solution found by Method 1, though in this case Method 1 was simpler. ■

When the indicial equation has either equal roots or roots differing by an integer, and only the leading terms (the most significant ones) are required in the second linearly independent solution $y_2(x)$, the reduction of order method is often the simplest one to use. This approach is illustrated in the following example, and it is typical of how best to proceed when the integrand in result (35) of Section 6.3 involves a quotient of polynomials.

EXAMPLE 8.11 Find the solution of

$$x^2y'' + (x^3 - x)y' + y = 0$$

in some interval $0 < x < d$.

Solution The equation has a singular point at the origin, and when it is written in standard form, we find that $P(x) = x - 1/x$ and $Q(x) = 1/x^2$. Thus,

$$p_0 = \lim_{x\to 0} xP(x) = -1 \quad \text{and} \quad q_0 = \lim_{x\to 0} x^2Q(x) = 1,$$

so the origin is a regular singular point and the indicial equation is

$$c(1-c) - c + 1 = 0 \quad \text{or} \quad (c-1)^2 = 0,$$

with the double root $c = 1$.

Making the substitution $y(x) = \sum_{n=0}^{\infty} a_nx^{n+c}$ in the differential equation gives

$$\sum_{n=0}^{\infty}(n+c)(n+c-1)a_nx^{n+c} + \sum_{n=0}^{\infty}(n+c)a_nx^{n+c+2} - \sum_{n=0}^{\infty}(n+c)a_nx^{n+c}$$
$$+ \sum_{n=0}^{\infty}a_nx^{n+c} = 0.$$

A shift of the summation index brings this to the form

$$(c^2-2c+1)x^c + c^2x^{c+1} + \sum_{n=2}^{\infty}(n+c)(n+c-1)a_nx^{n+c} + \sum_{n=2}^{\infty}(n+c-2)a_{n-2}x^{n+c}$$
$$-\sum_{n=2}^{\infty}(n+c)a_nx^{n+c} + \sum_{n=2}^{\infty}a_nx^{n+c} = 0,$$

and after combination of the summations this becomes

$$(c^2-2c+1)a_0x^c + c^2a_1x^{c+1} + \sum_{n=2}^{\infty}\{[(n+c)(n+c-2)+1]a_n$$
$$+ (n+c-2)a_{n-2}\}x^{n+c} = 0.$$

Equating the coefficient of x^c to zero gives the indicial equation with the double root $c = 1$, and equating the coefficient of x^{c+1} to zero shows that $a_1 = 0$, because $c = 1$. Equating the coefficient of x^{n+c} to zero leads to the recurrence relation

$$[(n+c)(n+c-2)+1]a_n + (n+c-2)a_{n-2} = 0 \quad \text{for } n \geq 2.$$

Setting $c = 1$ in the recurrence relation, we have

$$a_n = -\frac{(n-1)}{n^2}a_{n-2},$$

but as $a_1 = 0$, it follows immediately that $a_n = 0$ for all odd n. As a result we have

$$a_2 = -\frac{1}{2^2}a_0, \quad a_4 = -\frac{3}{4^2}a_2 = \frac{3}{2^2 \cdot 4^2}a_0, \quad a_6 = -\frac{5}{6^2}a_4 = -\frac{1\cdot3\cdot5}{2^2\cdot4^2\cdot6^2}a_0, \ldots,$$

so a fundamental solution is given by

$$y_1(x) = x\left(1 - \frac{1}{2^2}x^2 + \frac{1\cdot3}{2^2\cdot4^2}x^4 - \frac{1\cdot3\cdot5}{2^2\cdot4^2\cdot6^2}x^6 - \cdots\right),$$

or for $0 < x < d_1$, where d_1 is the radius of convergence of $y_1(x)$, by

$$y_1(x) = x - \frac{1}{2^2}x^3 + \frac{1\cdot3}{2^2\cdot4^2}x^5 - \frac{1\cdot3\cdot5}{2^2\cdot4^2\cdot6^2}x^7 + \cdots.$$

The reduction of order method in (35) of Section 6.3 shows that

$$y_2(x) = y_1(x)\int\frac{\exp\left[-\int P(x)dx\right]}{[y_1(x)]^2}dx,$$

but $\exp[-\int P(x)dx] = \exp(-x^2/2)$, so

$$y_2(x) = y_1(x)\int\frac{\exp(-x^2/2)}{[y_1(x)]^2}dx.$$

To find the leading terms in the expansion for $y_2(x)$ it is now necessary to replace $\exp(-x^2/2)$ and $[y_1(x)]^2$ by the first few terms of their series expansions and then to convert the integrand to a polynomial that can be integrated term by term. We have

$$y_2(x) = y_1(x)\int\frac{x\left(1-\frac{1}{2}x^2+\frac{1}{8}x^4-\frac{1}{48}x^6+\frac{1}{384}x^8-\cdots\right)}{x^2\left(1-\frac{1}{4}x^2+\frac{3}{64}x^4-\frac{5}{768}x^6+\frac{35}{49152}x^8-\cdots\right)^2}dx.$$

If the bracketed term in the denominator is now squared, the integral becomes

$$y_2(x) = y_1(x)\int \frac{\left(1 - \frac{1}{2}x^2 + \frac{1}{8}x^4 - \frac{1}{48}x^6 + \frac{1}{384}x^8 - \cdots\right)}{x\left(1 - \frac{1}{2}x^2 + \frac{5}{32}x^4 - \frac{7}{192}x^6 + \frac{169}{24576}x^8 - \cdots\right)}dx.$$

Division of the two polynomials using long division, or writing the numerator as

$$\frac{1}{x}\left(1 - \frac{1}{2}x^2 + \frac{1}{8}x^4 + \cdots\right)\left(1 - \frac{1}{2}x^2 + \frac{5}{32}x^4 - \cdots\right)^{-1}$$

and multiplying the bracketed terms after using the binomial theorem to expand the second bracket, converts the expression for $y_2(x)$ to

$$y_2(x) = y_1(x)\int \frac{1}{x}\left(1 - \frac{1}{32}x^4 + \frac{5}{8192}x^8 - \cdots\right)dx.$$

Integrating term by term, we find that for some $d_2 > 0$, the first few terms of the series solution $y_2(x)$ are

$$y_2(x) = y_1(x)\left[\ln\, x - \frac{1}{128}x^4 + \frac{5}{65536}x^8 + \cdots\right],$$

or

$$y_2(x) = y_1(x)\ln x + x\left(1 - \frac{1}{4}x^2 + \frac{3}{64}x^4 - \frac{5}{768}x^6 + \frac{35}{4915}x^8 - \cdots\right)\left(-\frac{1}{128}x^4 + \frac{5}{65536}x^8 + \cdots\right).$$

After multiplication of the two series we obtain

$$y_2(x) = y_1(x)\ln\, x - \left(\frac{1}{128}x^5 - \frac{59}{65536}x^9 + \cdots\right)$$

in some interval of the form $0 < x < d_2$, where d_2 is the radius of convergence of the bracketed series. The general solution is thus

$$y(x) = C_1 y_1(x) + C_2 y_2(x), \quad \text{for } 0 < x < d,$$

where C_1 and C_2 are arbitrary constants and $d = \min\{d_1, d_2\}$.

When using this approach it is important to ensure that sufficient terms are retained in the intermediate calculations involving the polynomials for the final result to be accurate to the required power of x. ■

Case (d) Complex Conjugate Roots

EXAMPLE 8.12 Find the solution of the Cauchy–Euler equation

$$x^2 y''(x) - xy'(x) + 10y(x) = 0$$

in some interval $0 < x < d$.

Solution This equation has a singular point at the origin, and when expressed in standard form $P(x) = -1/x$ and $Q(x) = 10/x^2$. We have

$$\lim_{x\to 0} xP(x) = -1 \quad \text{and} \quad \lim_{x\to 0} x^2 Q(x) = 10,$$

so the origin is a regular singular point. From (30) the indicial equation is seen to be

$$c^2 - 2c + 10 = 0$$

with the complex conjugate roots $c = 1 \pm 3i$. Substituting

$$y(x) = \sum_{n=0}^{\infty} a_n x^{n+c}$$

into the differential equation leads to the result

$$\sum_{n=0}^{\infty}(n+c)(n+c-1)a_n x^{n+c} - \sum_{n=0}^{\infty}(n+c)a_n x^{n+c} + \sum_{n=0}^{\infty} 10a_n x^{n+c} = 0.$$

After terms are collected under a single summation sign, this becomes

$$\sum_{n=0}^{\infty}[(n+c)(n+c-2)+10]a_n x^{n+c} = 0.$$

Equating to zero the coefficient of x^c, corresponding to $n = 0$, gives

$$(c^2 - 2c + 10)a_0 = 0,$$

but by hypothesis $a_0 \neq 0$, so this simply yields the indicial equation. Equating to zero the coefficient of x^{n+c} for $n = 1, 2, \ldots$ gives

$$(n+c)(n+c+10)a_n = 0,$$

but as $c = 1 \pm 3i$, the factor $(n+c)(n+c+10) \neq 0$ for any value of n, so it follows that $a_n = 0$ for $n = 1, 2, \ldots$. Thus, from Theorem 8.2(d), it follows that two linearly independent solutions of the differential equation are obtained by taking the real and imaginary parts of

$$\begin{aligned} y(x) &= a_0 x^{1+3i} = a_0 x \exp\{\ln x^{3i}\} = a_0 x \exp\{3i \ln x\} \\ &= a_0 x\{\cos(3\ln x) + i\sin(3\ln x)\}. \end{aligned}$$

Setting the arbitrary constant $a_0 = 1$ and taking the real and imaginary parts of this last result shows that two linearly independent solutions are

$$y_1(x) = x\cos\{3\ln x\} \quad \text{and} \quad y_2(x) = x\sin\{3\ln x\},$$

each of which is defined for $x > 0$. These solutions form a basis for the solution of the differential equation whose general solution is

$$y(x) = C_1 x\cos\{3\ln x\} + C_2 x\sin\{3\ln x\}, \quad \text{for } x > 0,$$

where C_1 and C_2 are arbitrary constants. ■

More information about singular points and the Frobenius method can be found in references [3.3] to [3.6].

Summary

This section showed how the power series solutions considered previously must be modified if solutions are to be obtained in the form of expansions about regular singular points. The method due to Frobenius for obtaining such solutions was then developed systematically and illustrated by examples, with particular attention being given to the various special cases that arise depending on the relationship that exists between the roots of the indicial equation.

EXERCISES 8.4

In Exercises 1 and 2, shift the summation indices to combine the given expressions into the sum of a finite number of terms and a single summation.

1. (a) $2\sum_{n=0}^{\infty} a_n x^{n+c} + (1+x)\sum_{n=0}^{\infty} a_n x^{n+c-2}$.

(b) $3\sum_{n=0}^{\infty} a_n x^{n+c} + 2x^2\sum_{n=0}^{\infty} a_n x^{n+c-1}$.

2. (a) $(x-x^3)\sum_{n=0}^{\infty} a_n x^{n+c} + 3\sum_{n=0}^{\infty} a_n x^{n+c-1}$.

(b) $(x^2-x)\sum_{n=0}^{\infty} a_n x^{n+c} + 2\sum_{n=0}^{\infty} a_n x^{n+c-2}$.

In Exercises 3 through 6, use long division and multiplication of series to find the first four terms of the given expressions.

3. (a) $\dfrac{1}{\sum_{n=0}^{\infty}(-1)^n x^n/(n+1)}$.

(b) $(1 - x/2 + x^2/4 - x^3/8 + x^4/16 - x^5/32 + \cdots)\exp(x)$.

(c) $(1 - x/2 + x^2/3 - x^3/4 + x^4/5 - \cdots)(1 - x + x^2/2 - x^3/3 + x^4/4 - \cdots)$.

4. (a) $(1+2x+x^2)/(3-x+2x^4)$.

(b) $\left(\sum_{n=1}^{\infty} \dfrac{x^n}{n^2}\right)\left(\sum_{n=1}^{\infty} \dfrac{(-1)^{n+1}x^n}{(n+1)}\right)$.

5. (a) $\displaystyle\int \frac{1}{x}\left(\frac{1-3x+x^2}{2-\exp(x)}\right)dx$. **(b)** $\displaystyle\int \frac{\exp x}{(x+x^2+x^3)}dx$.

6. (a) $\displaystyle\int \frac{1}{x^2}\frac{(1+2x-x^2)}{(1+x+2x^3)}dx$. **(b)** $\displaystyle\int \frac{1}{x}\frac{\exp(-x)}{(1-2x+2x^2)}dx$.

In Exercises 7 through 26, find two linearly independent solutions for $x > 0$, and determine at least the first four leading terms in the second solution $y_2(x)$.

7. $4x^2y'' + 2xy' + (x-2)y = 0$.
8. $3x^2y'' - xy' + (x+1)y = 0$.
9. $2x^2y'' + xy' - (2x+1)y = 0$.
10. $2x^2y'' + xy' - (3x+1)y = 0$.
11. $(x^2-1)y'' + 2xy' + y = 0$.
12. $2x^2y'' + 2xy' + (x^2-2)y = 0$.
13. $x(1-x)y'' + (1-x)y' - y = 0$.
14. $2x^2y'' - 2xy' + (x^2+2)y = 0$.
15. $x^2y'' + (2x^2-x)y' + y = 0$.
16. $x^2y'' + 2(x^2-x)y' + 2y = 0$.
17. $x^2y'' + (x^2-2x)y' + 2y = 0$.
18. $x^2y'' - xy' + (x^2+1)y = 0$.
19. $16x^2y'' + 8xy' + (16x+1)y = 0$.
20. $2x^2y'' + 2xy' + (x-2)y = 0$.
21. $x^2y'' + (x^2-x)y' - 3y = 0$.
22. $4x^2y'' - 2x^2y' + (2x+1)y = 0$.
23. $x^2y'' + (x^2+x)y' - 4y = 0$.
24. $9x^2y'' - 6xy' + 2y = 0$.
25. $x^2y'' - 4xy' + 20y = 0$.
26. $4x^2y'' + 8xy' + 5y = 0$.

27. By shifting the critical point to the origin, find two linearly independent solutions of the following equation in an interval of the form $0 < x+1 < d$:

$$2(x+1)y'' + y' - (x+1)y = 0.$$

28. By shifting the critical point to the origin, find two linearly independent solutions of the following equation in an interval of the form $0 < x-2 < d$:

$$(x-2)^2y'' - (x-2)y' + (x^2-4x+5)y = 0.$$

8.5 The Gamma Function Revisited

more about the Gamma function

The function $\Gamma(x)$, called the **gamma function**, was introduced in (4) of Section 7.1 in connection with the Laplace transform of t^a when a is not an integer, and it was defined in terms of the improper integral

$$\Gamma(x) = \int_0^{\infty} e^{-t}t^{x-1}\,dt \quad \text{for } x > 0. \tag{32}$$

a fundamental result

It was shown that $\Gamma(x)$ satisfies the recurrence relation

$$\Gamma(x+1) = x\Gamma(x) \quad \text{for } x > 0, \tag{33}$$

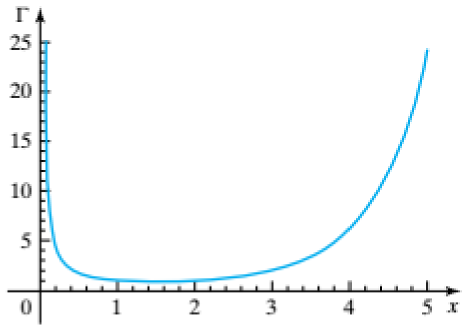


FIGURE 8.4 The function $\Gamma(x)$ in the interval $0 < x < 5$.

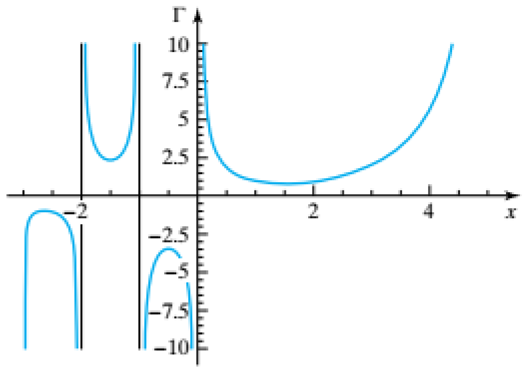


FIGURE 8.5 The function $\Gamma(x)$ in the interval $-3 < x < 4$.

and that when x is a positive integer n the gamma function reduces to

$$\Gamma(n+1) = n!. \tag{34}$$

Thus, for any real $x > 0$, the function $\Gamma(x)$ interpolates continuously between successive values of $n!$, and so generalizes the factorial function to nonintegral values of n. For obvious reasons the gamma function is sometimes called the **factorial function**. Figure 8.4 shows a graph of $\Gamma(x)$ in the interval $0 < x < 5$.

The gamma function can be extended to $x < 0$ for $x \neq -1, -2, \dots,$ at which point it becomes infinite. A graph of $\Gamma(x)$ in the interval $-3 < x < 4$ is shown in Fig. 8.5.

The value of $\Gamma(1/2)$ is often needed, and it can be found by means of the following method in which the integral defining $\Gamma(1/2)$ is squared and converted to a double integral that is easily evaluated. If the method used is unfamiliar the details can be omitted, though the result given in (35) is useful and should be remembered.

From (32) we have

$$[\Gamma(1/2)]^2 = \left(\int_0^\infty u^{-1/2}e^{-u}du\right)\left(\int_0^\infty v^{-1/2}e^{-v}dv\right),$$

where the two dummy variables u and v have been introduced to avoid confusion when the product of integrals is combined.

Writing $u = x^2$ and $v = y^2$ allows this product of integrals to be written as

$$[\Gamma(1/2)]^2 = \left(\int_0^\infty 2e^{-x^2}dx\right)\left(\int_0^\infty 2e^{-y^2}dy\right) = 4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}dxdy.$$

As the integral in terms of cartesian coordinates is only evaluated over the first quadrant, changing to the polar coordinates (r, θ) by setting $x = r\cos\theta$, $y = r\sin\theta$, and using the result $r^2 = x^2 + y^2$ reduces this last integral to

$$[\Gamma(1/2)]^2 = \lim_{\rho\to\infty} 4\int_0^{\pi/2} d\theta \int_0^\rho e^{-r^2} r\,dr = 4\cdot(\pi/2)\lim_{\rho\to\infty}\left[-\frac{1}{2}e^{-r^2}\right]_0^\rho = \pi.$$

a useful special case

Taking the square root shows that

$$\Gamma(1/2) = \sqrt{\pi}. \tag{35}$$

When x is a multiple of $1/2$, repeated use of recurrence relation (33) combined with result (35) allows $\Gamma(x)$ to be simplified, as illustrated in the following example.

EXAMPLE 8.13 Find (a) $\Gamma(7/2)$ and (b) $\Gamma(-3/2)$.

Solution

(a) From (33) it follows that

$$\Gamma\left(\frac{7}{2}\right) = \frac{5}{2}\Gamma\left(\frac{5}{2}\right) = \frac{5}{2}\cdot\frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{15}{8}\sqrt{\pi}.$$

(b) Setting $x = -3/2$ in (33) gives

$$\left(-\frac{3}{2}\right)\Gamma\left(-\frac{3}{2}\right) = \Gamma\left(-\frac{1}{2}\right),$$

whereas setting $x = -1/2$ in (33) gives

$$\left(-\frac{1}{2}\right)\Gamma\left(-\frac{1}{2}\right) = \Gamma(1/2) = \sqrt{\pi}.$$

So, combining these two results, we find that

$$\Gamma\left(-\frac{3}{2}\right) = \left(-\frac{2}{3}\right)\left(-\frac{2}{1}\right)\Gamma(1/2) = \frac{4}{3}\sqrt{\pi}.$$

The reason for this re-examination of the gamma function is because it enables the coefficients of a series expansion to be expressed in a concise form. For example, it follows directly from (34) that the binomial coefficient

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} = \frac{\Gamma(n+1)}{\Gamma(m+1)\Gamma(n-m+1)}. \tag{36}$$

Expressing a binomial coefficient with integer entries in terms of the gamma function offers no particular advantage over the use of factorials, but the preceding result generalizes to the more useful result

$$\binom{\alpha}{m} = \frac{\Gamma(\alpha+1)}{\Gamma(m+1)\Gamma(\alpha-m+1)} \tag{37}$$

when α is any nonnegative real number (not necessarily an integer). This expression is often useful when performing numerical calculations.

As another example of the use of (33) we notice that we can write

$$a(a+1)(a+2)\dots(a+n) = \frac{\Gamma(a+n+1)}{\Gamma(a)}, \tag{38}$$

where n is a positive integer and the real number $a > 0$. Thus, for example, in terms of the gamma function the following product becomes

$$\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\left(\frac{7}{2}\right) = \frac{\Gamma(\frac{1}{2}+3+1)}{\Gamma(\frac{1}{2})} = \frac{\Gamma(\frac{9}{2})}{\Gamma(\frac{1}{2})}.$$

Result (38) generalizes further to provide a concise representation of the product of $n+1$ factors $c(c+d)(c+2d)\ldots(c+nd)$. By writing the product as

$$c(c+d)(c+2d)\ldots(c+nd) = d^{n+1}\left(\frac{c}{d}\right)\left(\frac{c}{d}+1\right)\left(\frac{c}{d}+2\right)\cdots\left(\frac{c}{d}+n\right),$$

and then setting $a = c/d$ in (38), we arrive at the useful result

$$c(c+d)(c+2d)\ldots(c+nd) = d^{n+1}\frac{\Gamma\left(\frac{c}{d}+n+1\right)}{\Gamma\left(\frac{c}{d}\right)}. \tag{39}$$

EXAMPLE 8.14

The nth coefficient of a series is given by

$$a_n = \frac{1\cdot 5\cdot 9\cdot 13\ldots(4n+1)}{2^n}.$$

Express a_n in terms of the gamma function.

Solution Comparing the numerator of a_n with result (39) shows that it contains $n+1$ factors, and in the notation of (39) we have $c = 1$ and $d = 4$. Thus,

$$1\cdot 5\cdot 9\cdot 13\ldots(4n+1) = 4^{n+1}\frac{\Gamma\left(n+\frac{5}{4}\right)}{\Gamma\left(\frac{1}{4}\right)},$$

so dividing by 2^n we find that

$$a_n = 4^{n+1}\frac{\Gamma\left(n+\frac{5}{4}\right)}{2^n\Gamma\left(\frac{1}{4}\right)} = 2^{n+2}\frac{\Gamma\left(n+\frac{5}{4}\right)}{\Gamma\left(\frac{1}{4}\right)}.$$

Two special products of this type arise when working with series as, for example, occurs in the case of Legendre polynomials. These products involve either the product of consecutive pairs of odd numbers or the product of consecutive pairs of even numbers. Although these products can be expressed in terms of the gamma function, a convenient and concise **double factorial** notation is used. We define the double factorial !! as follows:

the double factorial

$$1\cdot 3\cdot 5\cdots(2n+1) = (2n+1)!! \quad \text{and} \quad 2\cdot 4\cdot 6\cdots(2n) = (2n)!!. \tag{40}$$

Alternative expressions for these double factorials in terms of the usual factorial function are

$$(2n+1)!! = \frac{(2n+1)!}{2^n n!} \quad \text{and} \quad (2n)!! = 2^n n!. \tag{41}$$

The following relationship connecting gamma functions is sometimes useful:

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}. \tag{42}$$

However, this result will not be proved here as it requires the techniques of complex integration.

the beta function

In passing, we mention a function $B(x, y)$ called the *beta function* that is related to the gamma function. The **beta function**, which has applications in statistics and elsewhere, is defined as the integral

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt \quad \text{with} \quad x > 0, y > 0. \tag{43}$$

The following are the most important properties of the beta function:

Symmetry:

$$B(x, y) = B(y, x) \tag{44}$$

relating gamma and beta functions

Connection with the gamma function:

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{45}$$

Relationship between beta functions:

$$B(x, y) = \left(\frac{y-1}{x+y-1}\right) B(x, y-1) = \left(\frac{x+y}{y}\right) B(x, y+1), \tag{46}$$

Special values:

$$B(1, 1) = 1 \quad \text{and} \quad B(1/2, 1/2) = \pi. \tag{47}$$

Outline proofs of results (42) to (44) will be found in the harder exercises at the end of this section.

The gamma function in the complex plane is discussed in reference [6.7], and general information about the gamma function and related functions is contained in Chapter 6 of reference [G.1] and Chapter 11 of reference [G.3].

Summary

The gamma function that was introduced earlier was seen to provide a natural extension to arbitrary values of x of the factorial function $n!$, where n is an integer. In this section the gamma function was examined in greater detail and some useful values were derived in terms of π. The beta function was then defined and related to the gamma function.

EXERCISES 8.5

1. Express $\Gamma(5/2), \Gamma(-5/2)$, and $\Gamma(9/2)$ in terms of $\sqrt{\pi}$.
2. Express $\Gamma(-9/2), \Gamma(11/2)$, and $\Gamma(-11/2)$ in terms of $\sqrt{\pi}$.
3. Express $\Gamma(5/4), \Gamma(-5/4)$, and $\Gamma(7/4)$ in terms of either $\Gamma(1/4)$ or $\Gamma(-1/4)$.
4. Express $\Gamma(-7/4), \Gamma(9/4)$, and $\Gamma(3/4)$ in terms of either $\Gamma(1/4)$ or $\Gamma(-1/4)$.
5. Express the product $6 \cdot 11 \cdot 16 \cdot 21 \ldots (5n+6)$ in terms of the gamma function.
6. Express the product $1 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \ldots .(2n+1)$ in terms of the gamma function.

7. Express the product $5 \cdot 8 \cdot 11 \cdot 14 \ldots (3n+5)$ in terms of the gamma function.
8. Express the product $4 \cdot 8 \cdot 12 \cdot 16 \ldots (4n+4)$ in terms of the gamma function.
9. Show that
$$\Gamma\left(\frac{1}{2} - n\right) = \frac{(-1)^n \sqrt{\pi}}{(n-\frac{1}{2})(n-\frac{3}{2})(n-\frac{5}{2})\cdots(\frac{1}{2})}.$$
10. Show that
$$\Gamma\left(n + \frac{1}{2}\right) = \left(n - \frac{1}{2}\right)\left(n - \frac{3}{2}\right)\cdots\left(\frac{1}{2}\right)\sqrt{\pi}.$$

The following slightly harder exercises provide more information about the gamma function.

11.* Use the result $\Gamma(n+\frac{1}{2}) = (n-\frac{1}{2})\Gamma(n-\frac{1}{2})$ with the result of Exercise 9 to show that
$$\Gamma(2n) = \frac{2^{2n-1}\Gamma(n)\Gamma\left(n+\frac{1}{2}\right)}{\sqrt{\pi}}.$$
12.* Show that $\Gamma(x) = \int_0^1 (\ln \frac{1}{u})^{x-1} du$ for $x > 0$.
13.* Show that $\Gamma(x) = 2\int_0^\infty e^{-u^2} u^{2x-1} du$ for $x > 0$.
14.* The function $\psi(x)$, called the **psi function** or the **digamma function**, is defined as
$$\psi(x) = \frac{d}{dx}[\ln \Gamma(x)].$$
Show that
$$\psi(x+1) = \psi(x) + \frac{1}{x} \quad \text{for } x > 0.$$
15.* Use the result of Exercise 14 to show that
$$\psi(x+n) = \psi(x) + \sum_{k=0}^{n-1} \frac{1}{x+k} \text{ where } n > 1 \text{ is an integer.}$$
16.* By making the variable change $u = 1 - t$ in the integral defining $B(x, y)$, show that $B(x, y) = B(y, x)$.
17.* Integrate $B(x, y)$ by parts to obtain the result of (46) that
$$B(x, y) = \left(\frac{y-1}{x+y-1}\right) B(x, y-1),$$
and use this result to obtain the second result of (46).
18.* Use the result of Exercise 17 to show that if m and n are integers,
$$B(m, n) = \frac{(m-n)!(n-1)!}{(m+m-1)!} = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)},$$
and so
$$B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

8.6 Bessel Function of the First Kind $J_n(x)$

Bessel's equation

In standard form, **Bessel's equation** is written
$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - \nu^2)y = 0, \tag{48}$$
where $\nu \geq 0$ is a real number. Another useful form of Bessel's equation that often arises in applications is
$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (\lambda^2x^2 - \nu^2)y = 0. \tag{49}$$
This form of the equation is obtained from (48) by first making the change of variable $x = \lambda u$, and then replacing u by x.

When developing the properties of Bessel functions in this section the standard form of the equation given in (48) will be used. Applications of Bessel functions to partial differential equations are made in Chapter 18.

Bessel's equation has a singularity at the origin, and using the notation of Section 8.4 with $P(x) = 1/x$ and $Q(x) = (x^2 - \nu^2)/x^2$, we find that
$$p_0 = \lim_{x\to 0} xP(x) = 1 \quad \text{and} \quad q_0 = \lim_{x\to 0} x^2Q(x) = -\nu^2,$$
showing that the origin is a regular singular point.

The indicial equation is seen to be

$$c^2 - \nu^2 = 0, \tag{50}$$

so the roots $c_1 = \nu$ and $c_2 = -\nu$ are distinct when $\nu \neq 0$, and there is a repeated zero root when $\nu = 0$. Thus, when $\nu = 0$, the second Frobenius solution will contain a logarithmic term, whereas when $c_1 - c_2$ is an integer the second Frobenius solution may or may not contain a logarithmic term. When $c_1 - c_2 \neq 0$ is not an integer, neither of the two linearly independent Frobenius solutions contains a logarithmic term.

Substituting $y(x) = \sum_{r=0}^{\infty} a_r x^{r+c}$ into (48) gives

$$\sum_{r=0}^{\infty}(r+c)(r+c-1)a_r x^{r+c} + \sum_{r=0}^{\infty}(r+c)a_r x^{r+c} + \sum_{r=0}^{\infty} a_r x^{r+c+2} - \nu^2 \sum_{r=0}^{\infty} a_r x^{r+c} = 0.$$

Shifting the summation index in the third summation and collecting terms under a single summation leads to the result

$$(c^2 - \nu^2)a_0 x^c + [(c+1)^2 - \nu^2]a_1 x^{c+1} + \sum_{r=2}^{\infty}[(r+c+\nu)(r+c-\nu)a_r + a_{r-2}]x^{r+c} = 0.$$

Equating the coefficients of powers of x to zero shows the following:

Coefficient of x^c:

$$(c^2 - \nu^2)a_0 = 0 \quad \text{(the indicial equation, because } a_0 \neq 0\text{)}$$

Coefficient of x^{c+1}:

$$[(c+1)^2 - \nu^2]a_1 = 0 \quad \text{(a condition on } a_1\text{)}$$

Coefficient of x^{r+c}:

$$[(r+c)^2 - \nu^2]a_r + a_{r-2} = 0 \quad \text{(a recurrence relation)} \tag{51}$$

As $(c+1)^2 - \nu^2 \neq 0$, it follows from the second result that $a_1 = 0$, and then from the recurrence relation (51) that $a_r = 0$ for all odd r. As only even indices r are involved in the recurrence relation, we set $r = 2m$ with $m = 0, 1, \ldots,$ after which substituting $c = \nu$ in the recurrence relation reduces it to

$$a_{2m} = -\frac{1}{4m(m+\nu)}a_{2m-2}, \quad \text{for } m = 1, 2, \ldots. \tag{52}$$

As a_0 is arbitrary, we normalize the solution in the standard manner by setting

$$a_0 = \frac{1}{2^\nu \Gamma(1+\nu)},$$

after which the coefficients a_{2m} become

$$a_2 = -\frac{a_0}{2^2(1+\nu)} = -\frac{1}{2^{2+\nu}1!\Gamma(2+\nu)}, \quad a_4 = -\frac{a_2}{2^2 2(2+\nu)} = \frac{1}{2^{4+\nu}2!\Gamma(3+\nu)}, \ldots,$$

and, in general,

$$a_{2m} = -\frac{(-1)^m}{2^{2m+\nu}m!\Gamma(m+1+\nu)}, \quad \text{for } m = 1, 2, \ldots. \tag{53}$$

the Bessel function $J_\nu(x)$

Using this result in the first Frobenius solution, which hereafter will be denoted by $J_\nu(x)$ and called a **Bessel function of the first kind of order** ν, gives

$$J_\nu(x) = x^\nu \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+\nu} m! \Gamma(m+1+\nu)} \quad \text{for } x \geq 0. \tag{54}$$

When $x < 0$ the corresponding expression for $J_\nu(x)$ follows from the preceding result by reversing the sign of x in the series and replacing x^ν by $|x|^\nu$. The ratio test shows the series for $J_\nu(x)$ to be absolutely convergent for all x.

So far ν has been an arbitrary nonnegative number, but the standard convention is that when ν is an integer it is denoted by n. Using the result that when $\nu = n$ the gamma function $\Gamma(m+1+n) = (m+n)!$ allows $J_n(x)$ to be written in the simpler form

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+n}}{2^{2m+n} m!(m+n)!}, \quad \text{for } n = 0, 1, 2, \ldots. \tag{55}$$

It was because of this use of n that, to avoid confusion, the summation index in the series was chosen to be m. The two most important special cases of (55) are:

Bessel functions $J_0(x)$ and $J_1(x)$

Bessel function of the first kind of order zero:

$$J_0(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m}(m!)^2} = 1 - \frac{x^2}{2^2(1!)^2} + \frac{x^4}{2^4(2!)^2} - \frac{x^6}{2^6(3!)^2} + \cdots \tag{56}$$

Bessel function of the first kind of order 1:

$$J_1(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{2^{2m+1} m!(m+1)!} = \frac{x}{2} - \frac{x^3}{2^3 1! 2!} + \frac{x^5}{2^5 2! 3!} - \frac{x^7}{2^7 3! 4!} + \cdots. \tag{57}$$

Graphs of $J_0(x)$, $J_1(x)$, and $J_2(x)$ are shown in Fig. 8.6.

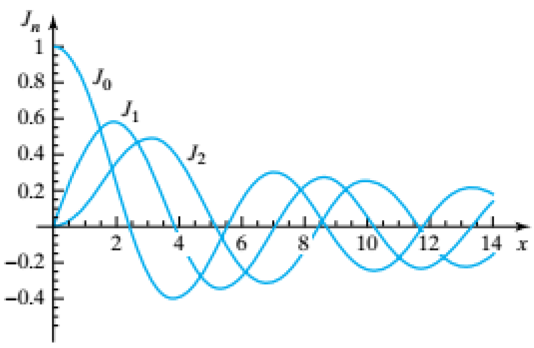


FIGURE 8.6 Graphs of the Bessel functions of the first kind $J_0(x)$, $J_1(x)$, and $J_2(x)$.

Having found $J_\nu(x)$, which is one solution of Bessel's equation (48), we must now find a second linearly independent solution in order to arrive at a basis for solutions of the equation, and hence to arive at the general solution. The nature of a second linearly independent solution will depend on the value of ν, and the simplest situation arises when ν is not an integer. In this case, because $c^2 = \nu^2$, a second linearly independent solution will follow from (54) by replacing ν by $-\nu$. Denoting this second solution by $J_{-\nu}(x)$ we find that

$$J_{-\nu}(x) = |x|^{-\nu} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m-\nu} m! \Gamma(m+1-\nu)} \quad \text{for } x \neq 0. \tag{58}$$

general solution of Bessel's equation

When ν is *not* an integer, the **general solution of Bessel's equation** (48) can be written

$$y(x) = C_1 J_\nu(x) + C_2 J_{-\nu}(x), \quad \text{for } x \neq 0, \tag{59}$$

with C_1 and C_2 arbitrary constants. The corresponding general solution of (49) is then

$$y(x) = C_1 J_\nu(\lambda x) + C_2 J_{-\nu}(\lambda x), \quad \text{for } x \neq 0. \tag{60}$$

The nature of the second linearly independent solution when $\nu = n$ will be considered later. In the meantime we will show that when $\nu = n$, the Bessel functions $J_n(x)$ and $J_{-n}(x)$ are linearly dependent. This is most easily seen by taking the limit of (58) as $\nu \to n$. Gamma functions with negative integer arguments are infinite, so the coefficients a_{2m} in which they occur will all vanish, causing the summation to start at the value $m = n$. Using the result $\Gamma(m+1-n) = (m-n)!$ then shows that the series for $J_{-n}(x)$ is

$$J_{-n}(x) = \sum_{m=n}^{\infty} \frac{(-1)^m x^{2m-n}}{2^{2m-n} m!(m-n)!},$$

and after a shift of the summation index this becomes

$$J_{-n}(x) = (-1)^n \sum_{m=n}^{\infty} \frac{(-1)^m x^{2m+n}}{2^{2m+n} m!(m+n)!}, \quad \text{for } n = 1, 2, \ldots. \tag{61}$$

A comparison of (55) and (61) shows that $J_{-n}(x)$ is a constant multiple of $J_n(x)$, so the two functions $J_n(x)$ and $J_{-n}(x)$ are linearly dependent. To be precise,

$$J_{-n}(x) = (-1)^n J_n(x) \quad \text{for } n = 1, 2, \ldots. \tag{62}$$

The absolute convergence of the series for $J_\nu(x)$ allows it to be differentiated term by term. Using this fact, and comparing of the derivative of the series for $J_0(x)$ with the series for $J_1(x)$, shows that

$$J_0'(x) = -J_1(x). \tag{63}$$

This result is the simplest example of the many relationships that exist between Bessel functions. The four most important results are the following:

relationships between derivatives and some recurrence relations

Relationships between derivatives of $J_\nu(x)$:

$$\frac{d}{dx}[x^\nu J_\nu(x)] = x^\nu J_{\nu-1}(x) \tag{64}$$

$$\frac{d}{dx}\left[x^{-\nu} J_\nu(x)\right] = -x^{-\nu} J_{\nu-1}(x) \tag{65}$$

Recurrence relations involving $J_\nu(x)$:

$$J_{\nu-1}(x) + J_{\nu+1}(x) = \frac{2\nu}{x} J_\nu(x) \tag{66}$$

$$J_{\nu-1}(x) - J_{\nu+1}(x) = 2J_\nu'(x) \tag{67}$$

We show next that these results are easily verified by substituting the series solution for $J_\nu(x)$ given in (54) into each relationship, though the direct derivation of these relationships is a more complicated matter. An indication of one way in which to arrive at these results without appealing to the series solution (54) is to be found in the set of exercises at the end of this section.

To establish (64) we start by multiplying the series (54) for $J_\nu(x)$ by x^ν to obtain

$$x^\nu J_\nu(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+2\nu}}{2^{2m+\nu} m! \Gamma(m+1+\nu)}.$$

Differentiating this result and removing a factor x^ν from the summation gives

$$\frac{d}{dx}[x^\nu J_\nu(x)] = x^\nu \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+\nu-1}}{2^{2m+\nu-1} \Gamma(m+\nu)},$$

but the series on the right-hand side is simply $J_{\nu-1}(x)$, so we have shown that

$$\frac{d}{dx}[x^\nu J_\nu(x)] = x^\nu J_{\nu-1}(x).$$

Result (65) is established in similar fashion by differentiating $x^{-\nu} J_\nu(x)$.

The recurrence relations can be obtained as follows. Carrying out the indicated differentiations and cancelling a factor x^ν in (64) and (65) gives

$$J_\nu'(x) = J_{\nu-1}(x) - \frac{\nu}{x} J_\nu(x) \tag{64$'$}$$

and

$$J_\nu'(x) = \frac{\nu}{x} J_\nu(x) - J_{\nu+1}(x). \tag{65$'$}$$

Results (66) and (67) now follow first by subtraction and then by addition of these two results.

Result (66) is useful because it relates $J_\nu(x)$ to $J_{\nu-1}(x)$ and $J_{\nu+1}(x)$, whereas (64) and (65) can be used to evaluate certain integrals involving $J_\nu(x)$, because by integrating (64) and (65) we obtain

$$\int x^\nu J_{\nu-1}(x)\mathrm{d}x = x^\nu J_\nu(x) + C \tag{68}$$

and

$$\int x^{-\nu} J_{\nu+1}(x)\mathrm{d}x = -x^{-\nu} J_\nu(x) + C. \tag{69}$$

EXAMPLE 8.15 Express $J_4(x)$ in terms of $J_0(x)$ and $J_1(x)$, and use the result to compute $J_4(6.2)$ given that $J_0(6.2) = 0.20175$ and $J_1(6.2) = -0.23292$.

Solution Rearranging (66) gives

$$J_{\nu+1}(x) = \frac{2\nu}{x} J_\nu(x) - J_{\nu-1}(x),$$

so setting $\nu = 3, 2,$ and 1 we have

$$J_4(x) = \frac{6}{x} J_3(x) - J_2(x), \quad J_3(x) = \frac{4}{x} J_2(x) - J_1(x), \text{ and } J_2(x) = \frac{2}{x} J_1(x) - J_0(x).$$

Eliminating $J_2(x)$ and $J_3(x)$ between these results gives the required expression

$$J_4(x) = \left(\frac{48}{x^3} - \frac{8}{x}\right) J_1(x) + \left(1 - \frac{24}{x^2}\right) J_0(x).$$

Setting $x = 6.2$ and substituting the given values of $J_0(6.2)$ and $J_1(6.2)$ shows that $J_4(6.2) = 0.32941$. ■

Numerical values of Bessel functions are extensively tabulated, and subroutines that enable their calculation for arbitrary values of their argument are found in most computer algebra packages. See the references at the end of the chapter for some of the most extensive tabulations of Bessel functions.

EXAMPLE 8.16 Evaluate

$$\int \left(x^2 + \frac{1}{x}\right) J_1(x) dx.$$

Solution We write the integral as the sum of integrals

$$\int \left(x^2 + \frac{1}{x}\right) J_1(x) dx = \int x^2 J_1(x) dx + \int x^{-1} J_1(x) dx$$

and consider each separately. Setting $\nu = 2$ in (64) shows that

$$\frac{d}{dx}[x^2 J_2(x)] = x^2 J_1(x),$$

so it follows at once that

$$\int x^2 J_1(x) dx = x^2 J_2(x) + C.$$

The second integral is a little harder and requires the use of integration by parts. Writing it as

$$\int x^{-1} J_1(x) dx = \int x^{-2}[x J_1(x)] dx,$$

and noticing from (63) with $\nu = 1$ that $[xJ_1(x)]' = xJ_0(x)$, we find that

$$\int x^{-1}J_1(x)dx = \int x^{-2}[xJ_1(x)]dx = -J_1(x) + \int x^{-1}xJ_0(x)dx,$$

and so

$$\int x^{-1}J_1(x)dx = -J_1(x) + \int J_0(x)dx.$$

No further simplification is possible because $\int J_0(x)\mathrm{d}x$ cannot be expressed in terms of simpler functions, though $\int_0^x J_0(u)du$ is available in tabular form and it is easily evaluated numerically on a computer. However, we will see later that $\int_0^\infty J_n(x)dx = 1$ for $n = 0, 1, 2, \ldots$.

EXAMPLE 8.17

Evaluate $\int x^3 J_0(x)\mathrm{d}x$.

Solution Writing the integrand as the product $x^3J_0(x) = x^2[xJ_0(x)]$ and using (64) with $\nu = 1$ gives

$$\int x^3 J_0(x)\mathrm{d}x = \int x^2[xJ_0(x)]\mathrm{d}x = \int x^2\frac{d}{dx}[xJ_1(x)]dx.$$

Integration by parts then gives

$$\int x^3 J_0(x)\mathrm{d}x = x^3J_1(x) - 2x^2J_2(x) + C.$$

It can be seen from Fig. 8.6 that the Bessel functions $J_0(x)$, $J_1(x)$, and $J_2(x)$ are oscillatory in nature and resemble damped sinusoids. The recurrence relation (66) implies that this same oscillatory property is true for all $J_n(x)$. Although these Bessel functions are not strictly periodic, in the sense that for any given n the zeros of $J_n(x)$ are not equally spaced along the x-axis, it can be shown that for fixed ν and large x the function $J_n(x)$ can be approximated by

$$J_\nu(x) \sim \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right), \tag{70}$$

where the symbol $\sim$ is to be read "is asymptotically equal to," with the understanding that the term *asymptotic* is used here in the technical sense and means that the ratio of the two sides of the expression tends to 1 as $x \to \infty$. This last result is an example of what is called an **asymptotic expansion** of the function $J_\nu(x)$, and asymptotic expansions have the property that the larger x becomes, the more accurate the asymptotic expansion becomes.

asymptotic expansion of $J_\nu(x)$

When the Bessel functions $J_\nu(x)$ are required in a computer program, the series solution (54) is used for small x, and different approximations are used for large x and in the intermediate region between small and large x. Corresponding approximations are used when the order ν of a Bessel function is large. The simplest approximation to $J_\nu(x)$ for small x, which follows from (54) by setting $m = 0$, is

$$J_\nu(x) \approx \frac{1}{\Gamma(1+\nu)}\left(\frac{x}{2}\right)^\nu. \tag{71}$$

The fact that the series for $J_\nu(x)$ is an alternating series means that the maximum magnitude of the error made when the series is truncated after n terms is the absolute

TABLE 8.1 Zeros $j_{n,r}$ of $J_n(x)$ for $n = 0, 1, 2, 3$

r	$j_{0,r}$	$j_{1,r}$	$j_{2,r}$	$j_{3,r}$
1	2.40482	3.83171	5.13162	6.38016
2	5.52007	7.01559	8.41724	9.76102
3	8.65372	10.17347	11.61984	13.01520
4	11.79153	13.32369	14.79595	16.22347
5	14.93091	16.47063	17.95982	19.40942
6	18.07106	19.61586	21.11700	22.58273

value of the $(n+1)$th term. So, if the series

$$J_0(x) = 1 - \frac{x^2}{2^2(1!)^2} + \frac{x^4}{2^4(2!)^2} - \frac{x^6}{2^6(3!)^2} + \cdots$$

is truncated after the term in x^4, the maximum error made is $|-x^6/[2^6(3!)^2]| = x^6/[2^6(3!)^2]$. Consequently, if $J_0(x)$ is approximated by

$$J_0(x) \approx 1 - \frac{x^2}{2^2(1!)^2} + \frac{x^4}{2^4(2!)^2},$$

then in the interval $0 \le x \le a$, the absolute value of the maximum error will not exceed $a^6/[2^6(3!)^2]$. When $J_\nu(x)$ is required to be accurate to a given number of decimal places in an interval $0 \le x \le a$, this simple estimate determines how many terms must be retained in the series approximation for $J_\nu(x)$.

zeros of Bessel functions $J_n(x)$

When using Bessel functions in applications, it is often necessary to know the location of the zeros of $J_n(x)$, so for future reference Table 8.1 lists the first six zeros of $J_n(x)$ for $n = 0, 1, 2, 3$. In the table the rth zero of $J_n(x)$ is denoted by $j_{n,r}$, where the first suffix indicates the order of the Bessel function and the second suffix the number of the zero. As $J_n(0) = 0$ for $n \ge 1$, the zeros $j_{1,r}$, $j_{2,r}$, and $j_{3,r}$ have been numbered so the first entry to appear in each column is the first nonvanishing zero of the function involved. Thus, although $J_1(0) = 0$, the first entry to appear in the column for $j_{1,r}$ is 3.83171, which it will be seen from Fig. 8.6 is the first nonvanishing zero of $J_1(x)$.

Bessel Functions $J_{\pm n/2}(x)$

The Bessel functions $J_{\pm n/2}(x)$ are particularly simple, despite the fact that the difference between the indices $c_1 = n/2$ and $c_2 = -n/2$ is an integer. The easiest way to find the form of $J_{\pm n/2}(x)$ is to use the reduction to standard form given in Lemma 6.1 of Section 6.3 to remove the first derivative term from Bessel's equation.

It follows from the lemma that the substitution $u = x^{1/2}y$ reduces Bessel's equation

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0$$

to the standard form for a second order equation

$$u'' + \left(1 - \frac{4\nu^2 - 1}{4x^2}\right)u = 0.$$

If we now consider the cases of $J_{1/2}(x)$ and $J_{-1/2}(x)$, corresponding to $\nu^2 = 1/4$, the differential equation simplifies to

$$u'' + u = 0,$$

with the general solution

$$u(x) = C_1 \sin x + C_2 \cos x.$$

As $y = x^{-1/2}u$, the general solution of Bessel's equation of order $\pm 1/2$ becomes

$$y(x) = C_1\sqrt{\frac{1}{x}}\sin x + C_2\sqrt{\frac{1}{x}}\cos x.$$

The two functions in the general solution for $y(x)$ are linearly independent, so we take for the solutions forming a basis for the differential equation with $\nu = \pm 1/2$ the functions $J_{1/2}(x)$ and $J_{-1/2}(x)$ given by

Bessel functions of fractional order

$$J_{1/2}(x) = C_1\sqrt{\frac{1}{x}}\sin x \quad \text{and} \quad J_{-1/2}(x) = C_2\sqrt{\frac{1}{x}}\cos x.$$

The constants C_1 and C_2 are arbitrary, but to make these results compatible with the normalization used for a_0 when developing the series solution for $J_\nu(x)$ we compare these expressions with the asymptotic formula (70), from which we see it is necessary to set $C_1 = C_2 = \sqrt{(2/\pi)}$, to obtain

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin x \quad \text{and} \quad J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\cos x. \tag{72}$$

Expressions for $J_{\pm n/2}(x)$ now follow by use of recurrence relation (66). Thus, for example, setting $\nu = 1/2$ in (66) gives

$$J_{3/2}(x) = \frac{1}{x}J_{1/2}(x) - J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\left(\frac{\sin x}{x} - \cos x\right), \tag{73}$$

and, similarly, setting $\nu = -1/2$ gives

$$J_{-3/2}(x) = -\sqrt{\frac{2}{\pi x}}\left(\sin x + \frac{\cos x}{x}\right). \tag{74}$$

We have shown that all Bessel functions $J_{\pm n/2}(x)$ with n an odd integer are expressible in terms of elementary functions. The derivation of $J_{\pm 1/2}(x)$ directly from series (54) forms an exercise in the set at the end of this section.

Friedrich Wilhelm Bessel (1784–1846)
A German mathematician who started his career as a clerk apprenticed to a mercantile office in Bremen where he remained for a number of years. Using published observations he calculated the orbit of Haley's comet and submitted his calculations to the astronomer H.W.M. Olbers who recognized his ability and, after recommending the work for publication, arranged for Bessel to become an assistant in the observatory in Lilienthal. His major mathematical contribution was the introduction, in a paper of 1824 devoted to planetary motions, of the class of transcendental functions now known as Bessel functions.

Summary

Bessel's equation was introduced and series solutions were obtained by the Frobenius method for the Bessel function $J_\nu(x)$ of the first kind of order ν. It was shown that Bessel functions of the first kind of fractional order $\pm n/2$, with n odd, could be expressed in terms of products of sines and cosines and $1/\sqrt{x}$.

EXERCISES 8.6

1. Write down the first six terms of the series expansion for $J_2(x)$.
2. Write down the first six terms of the series expansion for $J_3(x)$.
3. Derive result (65) by differentiating the product of $x^{-1/2}$ and the series for $J_\nu(x)$ given in (54).
4. Determine how many terms must be retained in the series for $J_0(x)$ for it to be accurate to four decimal places over the interval $0 \le x \le 4$.
5. Determine how many terms must be retained in the series for $J_0(x)$ for it to be accurate to four decimal places over the interval $0 \le x \le 2$.
6. Determine how many terms must be retained in the series for $J_0(x)$ for it to be accurate to six decimal places over the interval $0 \le x \le 1$.
7. Determine how many terms must be retained in the series for $J_0(x)$ for it to be accurate to six decimal places over the interval $0 \le x \le 2$.
8. Determine how many terms must be retained in the series for $J_1(x)$ for it to be accurate to four decimal places over the interval $0 \le x \le 2$.
9. Determine how many terms must be retained in the series for $J_1(x)$ for it to be accurate to four decimal places over the interval $0 \le x \le 3$.
10. Integrate the first four terms in the series for $J_0(x)$ term by term to obtain an approximation to
$$\int_0^x J_0(t)dt.$$
Estimate the maximum magnitude of the error when using the result in the interval $0 \le x \le a$.
11. Integrate the first four terms in the series for $J_1(x)$ term by term to obtain an approximation to
$$\int_0^x J_1(t)dt.$$
Estimate the maximum magnitude of the error when using the approximation in the interval $0 \le x \le a$. Integrate the integral analytically, and confirm that the analytical result and the approximation are in agreement.

The Bessel function $J_\nu(\lambda x)$ is a solution of $x^2y'' + xy' + (\lambda^2x^2 - \nu^2)y = 0$. Establish the following results by making the change of variable $x = \lambda X$ in results (64) to (67), and then replacing X by x.

12. $\dfrac{d}{dx}[x^\nu J_\nu(\lambda x)] = \lambda x^\nu J_{\nu-1}(\lambda x)$.
13. $\dfrac{d}{dx}[x^{-\nu} J_\nu(x)] = -\lambda x^{-\nu} J_{\nu+1}(\lambda x)$.
14. $\dfrac{d}{dx}[J_\nu(\lambda x)] = \lambda\, J_{\nu-1}(\lambda x) - \dfrac{\nu}{x}J_\nu(\lambda x)$.
15. $\dfrac{d}{dx}[J_\nu(\lambda x)] = -\lambda\, J_{\nu+1}(\lambda x) + \dfrac{\nu}{x}J_\nu(\lambda x)$.
16. $\dfrac{d}{dx}[J_\nu(\lambda x)] = \dfrac{\lambda}{2}[J_{\nu-1}(\lambda x) - J_{\nu+1}(\lambda x)]$.
17. $J_\nu(\lambda x) = \dfrac{\lambda x}{2\nu}[J_{\nu-1}(\lambda x) + J_{\nu+1}(\lambda x)]$.
18. Use (64)′ and (65)′ to show that
$$\frac{d}{dx}[xJ_\nu(x)J_{\nu+1}(x)] = x\left[J_\nu^2(x) - J_{\nu+1}^2(x)\right].$$
19. Show that $\lim_{x\to 0} J_0(x) = 1$, $\lim_{x\to 0} J_n(x) = 0$ for $n = 1, 2, \ldots$ and, $\lim_{x\to\infty} J_n(x) = 0$ for $n = 0, 1, \ldots,$ and prove that
$$\int_0^\infty J_1(x)dx = 1.$$
20. Use the results in Exercise 19 with (67) to show that
$$1 = \int_0^\infty J_1(x)ds = \int_0^\infty J_3(x)dx = \cdots$$
$$= \int_0^\infty J_{2n+1}(x)dx = \cdots \quad \text{for } n = 0, 1, \ldots.$$
21. In Section 7.3(d)(ii) it was shown that the Laplace transform of $J_0(x)$ was
$$\mathcal{L}\{J_0(x)\} = \frac{1}{(s^2+1)^{1/2}}.$$
Use this result to deduce the value of $\int_0^\infty J_0(x)dx$, and then use (67) together with the results of Exercise 20 to show that
$$1 = \int_0^\infty J_0(x)dx = \int_0^\infty J_1(x)dx = \int_0^\infty J_3(x)dx = \cdots$$
$$= \int_0^\infty J_n(x)dx = \cdots \quad \text{for } n = 0, 1, 2, \ldots.$$
22. Find (a) $\int x^3 J_2(x)\mathrm{d}x$ and (b) $\int x^{-3} J_4(x)\mathrm{d}x$.
23. Express $\int J_4(x)\mathrm{d}x$ in terms of $\int J_0(x)\mathrm{d}x$.

24. Express $\int J_5(x)\mathrm{d}x$ in terms of $J_0(x)$, $J_2(x)$, and $J_4(x)$.

25. Express $\int xJ_1(x)\mathrm{d}x$ in terms of $\int J_0(x)\mathrm{d}x$.

26. Express $\int x^2 J_0(x)\mathrm{d}x$ in terms of $\int J_0(x)\mathrm{d}x$.

The exercises that follow, some of which are slightly harder, provide background information about Bessel functions.

27.* By differentiating under the integral sign with respect to x, integrating by parts, and combining results using an elementary trigonometric identity, prove that

$$J_0(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin\theta)d\theta$$

is an integral representation of $J_0(x)$ by showing that it satisfies Bessel's equation of order zero

$$xJ_0'' + J_0' + xJ_0 = 0.$$

28.* The function $\exp[\frac{x}{2}(t - \frac{1}{t})]$ is the **generating function** for the Bessel functions $J_n(x)$, and it has the property that when it is expanded in powers of t (both positive and negative),

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

Thus, $J_n(x)$ is the coefficient of t^n in the expansion of the generating function in powers of t. Expand the exponential as the product of the series for $\exp[xt/2]$ and $\exp[-x/(2t)]$, and hence derive the first three terms of the series expansion of $J_0(x)$.

29.* Differentiate the generating function partially with respect to x and equate the coefficients of t^n on each side of the identity to prove that

$$2J_n'(x) = J_{n-1}(x) - J_{n+1}(x).$$

30.* Differentiate the generating function partially with respect to t and equate the coefficients of t^{n-1} on each side of the identity to prove that

$$\frac{2n}{x}J_n(x) = J_{n-1}(x) + J_{n+1}(x).$$

31.* Substitute $\nu = 1/2$ in (54) and (58), and hence show that $J_{1/2} = \sqrt{\frac{2}{\pi x}}\sin x$ and $J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\cos x$.

32.* Use (66) together with results (73) and (74) to show that

$$J_{5/2}(x) = \sqrt{\frac{2}{\pi x}}\left[\left(\frac{3}{x^2} - 1\right)\sin x - \frac{3}{x}\cos x\right]$$

$$J_{-5/2}(x) = \sqrt{\frac{2}{\pi x}}\left[\frac{3}{x}\sin x + \left(\frac{3}{x^2} - 1\right)\cos x\right]$$

$$J_{9/2}(x) = \sqrt{\frac{2}{\pi x}}\left[\left(\frac{105}{x^4} - \frac{45}{x} + 1\right)\sin x - \left(\frac{105}{x^3} - \frac{10}{x}\right)\cos x\right]$$

and

$$J_{-9/2}(x) = \sqrt{\frac{2}{\pi x}}\left[\left(\frac{105}{x^3} - \frac{10}{x}\right)\sin x + \left(\frac{105}{x^4} - \frac{45}{x^2} + 1\right)\cos x\right].$$

8.7 Bessel Functions of the Second Kind $Y_\nu(x)$

It was shown in the previous section that, with the exception of $\nu = 1/2$, the two Bessel functions $J_\nu(x)$ and $J_{-\nu}(x)$ of the first kind are only linearly independent solutions of Bessel's equation when the roots of the indicial equation differ by an integer. So it remains for us to find a second linearly independent solution when $\nu = n$ and $n = 0, 1, 2, \ldots$. We begin by considering the case $n = 0$, corresponding to the repeated root $\nu = 0$, when it follows from Theorem 8.2(b) that the form of solution to be expected in the case of Bessel's equation of order zero

Bessel functions of the second kind

$$xy'' + y' + xy = 0 \tag{75}$$

is

$$y_2(x) = J_0(x)\ln x + \sum_{r=0}^{\infty} b_r x^{r+1}. \tag{76}$$

Differentiation of (76) gives

$$y_2'(x) = J_0'(x)\ln x + \frac{J_0(x)}{x} + \sum_{r=0}^{\infty}(r+1)b_r x^r$$

and

$$y_2''(x) = J_0''(x)\ln x + \frac{2J_0'(x)}{x} - \frac{J_0(x)}{x^2} + \sum_{r=0}^{\infty}(r+1)rb_r x^{r-1}.$$

When these expressions are substituted into (75) the terms in $J_0(x)$ cancel, causing the equation to reduce to

$$[xJ_0''(x) + J_0'(x) + xJ_0(x)]\ln x + 2J_0'(x) + \sum_{r=0}^{\infty}(r+1)rb_r x^r$$
$$+ \sum_{r=0}^{\infty}(r+1)b_r x^r + \sum_{r=0}^{\infty}b_r x^{r+2} = 0.$$

The logarithmic term vanishes because $J_0(x)$ is a solution of (75), so the coefficients b_r are determined by the equation

$$2J_0'(x) + \sum_{r=0}^{\infty}(r+1)rb_r x^r + \sum_{r=0}^{\infty}(r+1)b_r x^r + \sum_{n=0}^{\infty}b_r x^{r+2} = 0.$$

To proceed further it is necessary to determine $J_0'(x)$, but this can be found by differentiating (56) in Section 8.6. After cancellation of a factor $2m$ from the numerator and denominator of the resulting expression, and noticing that the summation now starts from $m = 1$, it is found that

$$J_0'(x) = \sum_{m=1}^{\infty}\frac{(-1)^m x^{2m-1}}{2^{2m-1}(m-1)!m!}.$$

Combining this with the previous result gives

$$\sum_{m=1}^{\infty}\frac{(-1)^{m+1} x^{2m+1}}{2^{2m}(m+1)!m!} + \sum_{r=0}^{\infty}(r+1)rb_r x^r + \sum_{r=0}^{\infty}(r+1)b_r x^r + \sum_{r=0}^{\infty}b_r x^{r+2} = 0.$$

Shifting the summation index in the last term and combining the summations reduces this to

$$\sum_{m=1}^{\infty}\frac{(-1)^{m+1} x^{2m+1}}{2^{2m}(m+1)!m!} + b_0 + 4b_1 x + \sum_{r=2}^{\infty}\left\{(r+1)^2 b_r + b_{r-2}\right\}x^r = 0.$$

We now make use of the fact that terms may be rearranged in an absolutely convergent series in order to rewrite the last summation as a sum of even powers of x and a sum of odd powers of x before combining the results. The preceding equation then becomes

$$\sum_{m=1}^{\infty}\frac{(-1)^{m+1} x^{2m+1}}{2^{2m}(m+1)!m!} + b_0 + 4b_1 x + \sum_{m=1}^{\infty}\left\{(2m+1)^2 b_{2m} + b_{2m-2}\right\}x^{2m}$$
$$+ \sum_{m=2}^{\infty}[4m^2 b_{2m-1} + b_{2m-3}]x^{2m-1} = 0.$$

Next we equate the coefficient of each power of x to zero in the usual manner. As there is no constant term in the first summation, it follows that $b_0 = 0$. The recurrence relation in the second summation is $(2m+1)^2 b_{2m} + b_{2m-2} = 0$, so together with the result $b_0 = 0$ this implies that $b_{2m} = 0$ for $m = 0, 1, 2, \ldots$. Setting the summation involving even powers of x to zero brings the equation into the form

$$\sum_{m=1}^{\infty} \frac{(-1)^{m+1}\, x^{2m+1}}{2^{2m}(m+1)!m!} + 4b_1 x + \sum_{m=2}^{\infty} \left[4m^2 b_{2m-1} + b_{2m-3}\right] x^{2m-1} = 0.$$

We now equate to zero the coefficients of each remaining power of x, and proceeding in this manner it is not difficult to show that the general coefficient b_{2m-1} can be written

$$b_{2m-1} = \frac{(-1)^{m-1}}{2^{2m}(m!)^2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}\right), \quad \text{for } m = 1, 2, \ldots,$$

so the second linearly independent solution is

$$y_2(x) = J_0(x)\ln x + \sum_{m=1}^{\infty} \frac{(-1)^{m-1}\, x^{2m}}{2^{2m}(m!)^2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}\right). \tag{77}$$

Defining h_m as

$$h_m = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m} \tag{78}$$

allows $y_2(x)$ to be written in the more convenient form

$$y_2(x) = J_0(x)\ln x + \sum_{m=1}^{\infty} \frac{(-1)^{m-1}\, h_m x^{2m}}{2^{2m}(m!)^2}. \tag{79}$$

The series in (79) can be shown to converge, though as the logarithmic term becomes infinite at the origin, result (79) is only finite for $x > 0$.

As any linear combination of two linearly independent solutions of a differential equation is itself a solution, it proves to be convenient to take as the second solution of Bessel's equation of order zero the function $Y_0(x)$ defined as the linear combination

$$Y_0(x) = \frac{2}{\pi}[y_2(x) + (\gamma - \ln 2)J_0(x)], \tag{80}$$

where the constant γ, called the **Euler constant**, is defined as

$$\gamma = \lim_{m\to\infty}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m} - \ln m\right), \tag{81}$$

where $\gamma = 0.577\ 215\ 664\ 901\ldots$. This constant is also called the **Euler–Mascheroni constant**, and on occasion it is denoted by C and sometimes by $\ln \gamma$.

the Bessel functions $Y_0(x)$ and $Y_\nu(x)$

The function $Y_0(x)$, called the **Bessel function of the second kind of order zero**, is defined as

$$Y_0(x) = \frac{2}{\pi}\left[J_0(x)\left(\ln\frac{x}{2}+\gamma\right)+\sum_{m=1}^{\infty}\frac{(-1)^{m-1}h_m}{2^{2m}(m!)^2}x^{2m}\right]. \tag{82}$$

The reason for choosing this particular combination of functions in the definition of $Y_0(x)$ is because of its convenient properties as $x \to \infty$. The function $Y_0(x)$ is also called the **Neumann** or **Weber function** of order zero and denoted by $N_0(x)$.

Some authors make a distinction in what they call a Bessel function of the second kind, so there may be a difference between the Weber function $Y_n(x)$ and the Neumann function $N_n(x)$. Because of this, care must be exercised when using these functions in software packages.

Bessel functions of the second kind of integral order can be defined in similar fashion, but to make them compatible with the functions $J_{-\nu}(x)$ introduced in Section 8.6 the following definition is adopted:

$$Y_\nu(x) = \frac{1}{\sin\nu\pi}[J_\nu(x)\cos\nu\pi - J_{-\nu}(x)] \tag{83}$$

with

$$Y_n(x) = \lim_{\nu\to n} Y_\nu(x). \tag{84}$$

Using this last result it is possible to show that for integral values of ν the function $Y_n(x)$ is given by

$$\begin{aligned}Y_n(x) = \frac{2}{\pi}J_n(x)\left(\ln\frac{x}{2}+\gamma\right)+\frac{x^n}{\pi}\sum_{m=0}^{\infty}\frac{(-1)^{m-1}(h_m+h_{m+n})}{2^{2m+n}\,m!(m+n)!}x^{2m}\\ -\frac{1}{\pi x^n}\sum_{m=0}^{n-1}\frac{(n-m-1)!}{2^{2m-n}m!}x^{2m}\end{aligned} \tag{85}$$

where, by definition, $h_0 = 1$. It follows from this that the Bessel functions $Y_n(x)$ and $Y_{-n}(x)$ are linearly dependent, with

$$Y_{-n}(x) = (-1)^n Y_n(x).$$

Graphs of the first three Bessel functions of the second kind are shown in Fig. 8.7.

When x is small the following approximations are useful:

$$Y_0(x) \approx \frac{2}{\pi}\ln x \quad \text{and for } \nu > 0, \quad Y_\nu(x) \approx -\frac{\Gamma(\nu)}{\pi}\left(\frac{2}{x}\right)^\nu. \tag{86}$$

asymptotic form for $Y_\nu(x)$

For large x, however, the asymptotic approximation to $Y_\nu(x)$ is

$$Y_\nu(x) \sim \sqrt{\frac{2}{\pi x}}\sin\left[x-\left(\frac{2\nu+1}{4}\right)\pi\right]. \tag{87}$$

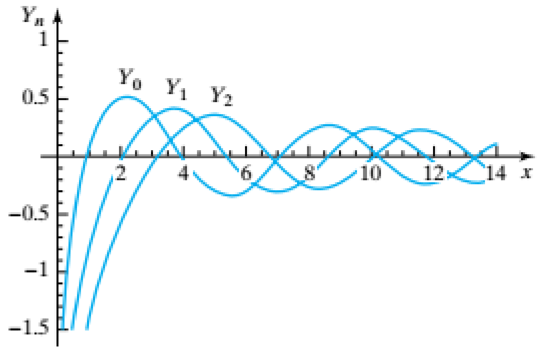


FIGURE 8.7 Bessel functions $Y_0(x)$, $Y_1(x)$, and $Y_2(x)$ of the second kind.

It follows from (86) and (87) that

$$\lim_{x\to 0} Y_\nu = -\infty \quad \text{and} \quad \lim_{x\to\infty} Y_\nu(x) = 0. \tag{88}$$

zeros of Bessel functions $Y_n(x)$

The zeros of $Y_n(x)$ are needed when working with Bessel functions, so the locations of the first six zeros of $Y_n(x)$ for $n = 0, 1, 2, 3$ are listed in Table 8.2. The rth zero of the Bessel function $Y_n(x)$ is denoted by $y_{n,r}$, so, for example, the second zero of $Y_1(x)$ is $y_{1,2} = 5.42968$.

It is a consequence of the definition of $Y_\nu(x)$ that for all ν the general solution of Bessel's equation in the standard form

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0 \tag{89}$$

is

$$y(x) = C_1 J_\nu(x) + C_2 Y_\nu(x). \tag{90}$$

Similarly, the general solution of Bessel's equation in the form

$$x^2y'' + xy' + (\lambda^2x^2 - \nu^2)y = 0 \tag{91}$$

is

$$y(x) = C_1 J_\nu(\lambda x) + C_2 Y_\nu(\lambda x). \tag{92}$$

TABLE 8.2 Zeros $y_{n,r}$ of $Y_n(x)$ for $n = 0, 1, 2, 3$

r	$y_{0,r}$	$y_{1,r}$	$y_{2,r}$	$y_{3,r}$
1	0.89358	2.19714	3.38424	4.52702
2	3.95786	5.42968	6.79381	8.09755
3	7.08605	8.59601	10.02348	11.39647
4	10.22235	11.74915	13.20999	14.62308
5	13.36110	14.89744	16.37897	17.81846
6	16.50092	18.04340	19.53904	20.99728

Many differential equations can be solved in terms of Bessel functions after a suitable transformation of the dependent variable. In particular, the equation

$$y'' + \left(\frac{1-2a}{x}\right)y' + \left[b^2c^2x^{2c-2} + \left(\frac{a^2 - \nu^2c^2}{x^2}\right)\right]y = 0 \tag{93}$$

can be shown to have the solution

$$y(x) = x^a Z_\nu(bx^c), \tag{94}$$

where a, b, and c are numbers and Z_ν is any linear combination of J_ν and Y_ν (see Exercise 16 at the end of this section).

The following is an application of Bessel functions to a simple physical problem. It illustrates how, in this case, the conditions of the problem only allow a Bessel function of the first kind to be retained in the solution. The problem, which is a classical one, can be stated as follows.

Find the radial temperature distribution $T(r)$ in a wire of circular cross-section with $0 \le r \le R$, when the electrical conductivity is σ, the thermal conductivity is K, and the wire carries a uniform current of density I amps per unit area of cross-section. Assume that the temperature at the center of the wire is T_0 and that the resistance of the wire varies linearly with the temperature as $\alpha T(r)$, with α a constant.

In order to formulate the problem in mathematical terms, we begin with the fact that the rate of heat generation in a unit volume of the wire is given by JI^2/σ heat units, where J is a physical constant (typically the number of calories in a joule). It follows from arguments given later in Chapter 18 that the equation determining the radial steady state temperature distribution is

$$K\frac{d^2T}{dr^2} + \frac{K}{r}\frac{dT}{dr} + \frac{\alpha JI^2}{\sigma}T = -\frac{JI^2}{\sigma},$$

where the last term on the left takes account of the linear variation of resistance with temperature, and the term on the right represents the heat generation due to the current.

When divided by K, this is seen to be Bessel's equation of order zero with a nonhomogeneous term $-JI^2/K\sigma$, and it is easily shown to have the general solution

$$T(r) = AJ_0\left(Ir\sqrt{\frac{\alpha J}{K\sigma}}\right) + BY_0\left(Ir\sqrt{\frac{\alpha J}{K\sigma}}\right) - \frac{1}{\alpha},$$

with A and B arbitrary constants. As the temperature must remain finite at the center of the wire, we must set $B = 0$ to remove the infinite value of Y_0 when $r = 0$. However, $T(0) = T_0$, so $A = T_0 + 1/\alpha$ and the required radial temperature distribution becomes

$$T(r) = \left(T_0 + \frac{1}{\alpha}\right)J_0\left(Ir\sqrt{\frac{\alpha J}{K\sigma}}\right) - \frac{1}{\alpha} \quad \text{for} \quad 0 \le r \le R. \qquad \blacksquare$$

Summary

It was seen in the previous section that when n is an integer $J_n(x)$ and $J_{-n}(x)$ are linearly dependent. This section has shown how a second linearly independent solution $Y_\nu(x)$ can be constructed that for all ν is linearly independent of $J_\nu(x)$, so the general solution of Bessel's equation can always be written $y(x) = AJ_\nu(x) + BY_\nu(x)$, where A and B are arbitrary constants. The function $Y_\nu(x)$ is called a Bessel function of the second kind of order ν.

EXERCISES 8.7

In Exercises 1 through 10, find the general solution of the differential equation.

1. $x^2y'' + xy' + (x^2 - 4)y = 0.$
2. $4x^2y'' + 4xy' + (4x^2 - 1)y = 0.$
3. $xy'' + y' + xy = 0.$
4. $xy'' + y' + \lambda^2 xy = 0.$
5. $xy'' + y' + 4x^3y = 0$; substitute $u = x^2$.
6. $x^2y'' + 3xy' + (x^2 + 1)y = 0$; substitute $y = u/x$.
7. $x^2y'' + xy' + 4(x^2 - 1)y = 0.$
8. $xy'' + y' + 9x^5y = 0$; substitute $u = x^3$.
9. $4x^2y'' + (16x^2 + 1)y = 0$; substitute $y = x^{1/2}u$.
10. $x^2y'' + 5xy' + (x^2 + 4)y = 0$; substitute $y = u/x^2$.

Use (93) and (94) to find the solution of the differential equations in Exercises 11 through 15.

11. $x^2y'' - xy' + (4x^4 - 3)y = 0.$
12. $xy'' - 3y' + xy = 0.$
13. $x^2y'' - xy' + (9x^2 + 1)y = 0.$
14. $x^2y'' - 5xy' + (16x^4 + 1)y = 0.$
15. $x^2y'' - 3xy' + (64x^8 - 8)y = 0.$
16. Verify that $y(x) = x^a Z_\nu(bx^c)$ is a solution of (93) by substituting for $y(x)$ in the differential equation and showing that this leads to the equation
$$X^2 Z_\nu''(X) + XZ'_\nu(X) + (X^2 - \nu^2)Z_\nu(X) = 0,$$
with $X = bx^c$. Hence, conclude that $Z_\nu(X)$ is either $J_\nu(X)$ or $Y_\nu(X)$, and so, because of the linearity of the equation, $Z_\nu(X) = C_1J_\nu(X) + C_2Y_\nu(X)$ must be a solution.
17. Use the substitution $y(x) = x^{-\nu}u(x)$ to convert the equation
$$x^2\frac{d^2y}{dt^2} + ax\frac{dy}{dx} + (1 + k^2x^2)y = 0,$$
in which a is a parameter, into an equation for $u(x)$. Find the values of a and ν that make the equation in $u(x)$ Bessel's equation of order zero. Use the result to find the general solution $y(x)$ that corresponds to this value of a.

8.8 Modified Bessel Functions $I_\nu(x)$ and $K_\nu(x)$

Replacing the independent variable x in Bessel's equation by ix changes the differential equation to

$$x^2y'' + xy' - (x^2 + \nu^2)y = 0, \tag{95}$$

Bessel's modified equation

called **Bessel's modified equation of order** ν.

It follows directly from Section 8.7 that Bessel's modified equation has two linearly independent complex solutions $J_\nu(ix)$ and $Y_\nu(ix)$. These solutions are not convenient to use, so the process of scaling and combining linearly independent solutions of a linear differential equation to form other solutions is used to produce two real linearly independent solutions denoted by $I_\nu(x)$ and $K_\nu(x)$. These are called, respectively, **modified Bessel functions of the first and second kinds of order** ν.

The modification of $J_\nu(ix)$ is straightforward, because from (54)

$$J_\nu(ix) = \sum_{m=0}^{\infty} \frac{(-1)^m(ix)^{2m+\nu}}{2^{2m+\nu}m!\Gamma(m+1+\nu)} = i^\nu \sum_{m=0}^{\infty} \frac{x^{2m+\nu}}{2^{2m+\nu}m!\Gamma(m+1+\nu)},$$

so the factor i^ν is removed and the **modified Bessel function of the first kind of order** ν is defined as the real function

$$I_\nu(x) = \sum_{m=0}^{\infty} \frac{x^{2m+\nu}}{2^{2m+\nu}\,m!\Gamma(m+1+\nu)}. \tag{96}$$

Unlike the series for $J_\nu(x)$, the series for $I_\nu(x)$ in (96) is no longer an alternating series, though it converges rapidly. As with ordinary Bessel functions, provided ν is not an integer, the general solution of Bessel's modified equation (95) can be written

the modified Bessel functions $I_\nu(x)$ and $K_\nu(x)$

$$y(x) = C_1 I_\nu(x) + C_2 I_{-\nu}(x). \tag{97}$$

However, rather than use $I_{-\nu}(x)$, in its place it is usual to introduce the real function $K_\nu(x)$ defined as the linear combination of real functions

$$K_\nu(x) = \left(\frac{\pi}{2}\right)\left(\frac{I_{-\nu}(x) - I_\nu(x)}{\sin \nu\pi}\right), \tag{98}$$

and to call $K_\nu(x)$ the **modified Bessel function of the second kind of order** ν. It can be seen from (98) that the functions $I_\nu(x)$ and $K_\nu(x)$ are linearly independent.

The definition of $K_\nu(x)$ can be extended to the case in which ν is an integer n by defining the function $K_n(x)$ as

$$K_n(x) = \lim_{\nu \to n} \left(\frac{\pi}{2}\right)\left(\frac{I_{-\nu}(x) - I_\nu(x)}{\sin \nu\pi}\right). \tag{99}$$

Because of this extension of the definition of $K_\nu(x)$, the general solution of Bessel's modified equation (95) can always be written in the form

$$y(x) = C_1 I_\nu(x) + C_2 K_\nu(x), \tag{100}$$

with no restriction placed on ν. The function $K_\nu(x)$ is also sometimes called the Kelvin function.

Similarly, when Bessel's modified equation is written in the form

$$x^2 y'' + xy' - (\lambda^2 x^2 + \nu^2)y = 0, \tag{101}$$

its general solution is given by

$$y(x) = C_1 I_\nu(\lambda x) + C_2 K_\nu(\lambda x), \tag{102}$$

with no restriction placed on ν.

This definition of $K_0(x)$ leads to the expansion

$$\begin{aligned} K_0(x) = &-\left[\ln \frac{x}{2} + \gamma\right] I_0(x) + \frac{x^2/4}{(1!)^2} + \left(1 + \frac{1}{2}\right)\frac{(x^2/4)^2}{(2!)^2} \\ &+ \left(1 + \frac{1}{2} + \frac{1}{3}\right)\frac{(x^2/4)^3}{(3!)^2} + \cdots, \end{aligned} \tag{103}$$

with similar though more complicated expansions for $K_n(x)$.

Graphs of $I_0(x)$ and $I_1(x)$ and of $K_0(x)$ and $K_1(x)$ are shown in Figs. 8.8 and 8.9, respectively.

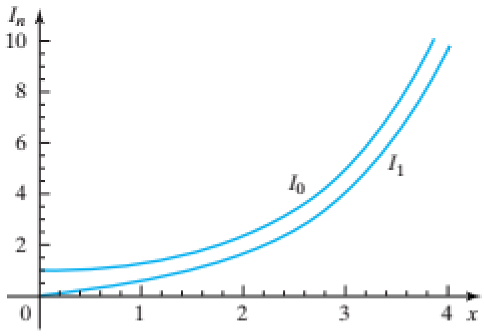


FIGURE 8.8 Graphs of $I_0(x)$ and $I_1(x)$.

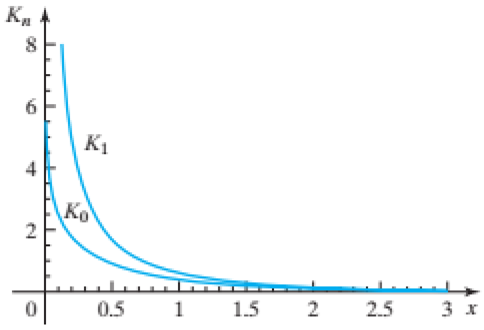


FIGURE 8.9 Graphs of $K_0(x)$ and $K_1(x)$.

The following are useful properties of $I_\nu(x)$ and $K_\nu(x)$:

$$I_0(0) = 1, \quad I_n(0) = 0 \quad \text{for } n = 1, 2, \ldots, \lim_{x \to 0} I_\nu(x) = 0,$$
$$K_n(0) = \infty, \quad \lim_{x \to \infty} K_n(x) = 0 \quad \text{for } n = 0, 1, 2, \ldots. \tag{104}$$

asymptotic expressions for modified Bessel functions

For small x

$$I_\nu(x) \sim \frac{1}{\Gamma(1+\nu)}\left(\frac{x}{2}\right)^\nu, \quad K_0(x) = -\ln x \quad \text{and}$$
$$K_\nu(x) \sim \frac{\Gamma(\nu)}{2}\left(\frac{2}{x}\right)^\nu \quad \text{for } \nu > 0, \tag{105}$$

whereas for large x

$$I_\nu(x) \approx \frac{1}{\sqrt{2\pi x}} e^x \quad \text{and} \quad K_\nu(x) \approx \sqrt{\frac{\pi}{2x}} e^{-x}. \tag{106}$$

Results involving Bessel functions of the first and second kinds, together with applications, are to be found in Chapter 5 of reference [3.7]. Chapters 9 to 11 of Reference [G.1] and Chapter 17 of reference [G.3] give general information about all types of Bessel functions. The standard encyclopedic work covering all aspects of Bessel functions is reference [3.17].

Summary

Modified Bessel functions were introduced, their series solutions were obtained, the general solution was expressed in terms of $I_\nu(x)$ and $K_\nu(x)$, and asymptotic representations were given.

EXERCISES 8.8

1. By differentiating the series for $I_0(x)$, show that $I_0'(x) = I_1(x)$.

2. Use the definition of $I_\nu(x)$ to show that

$$I_{\nu-1}(x) - I_{\nu+1}(x) = \frac{2\nu}{x} I_\nu(x) \quad \text{for } \nu \geq 1.$$

3. Use the definition of $I_\nu(x)$ to show that

$$I_{\nu-1}(x) + I_{\nu+1}(x) = 2I_\nu'(x) \quad \text{for } \nu \geq 1.$$

4. Use Lemma 6.1 of Section 6.3 to reduce Bessel's modified equation of order $\nu = 1/2$ to standard form, and hence show that

$$I_{1/2}(x) \text{ is proportional to } \frac{\sinh x}{\sqrt{x}}, \quad \text{and}$$

$$I_{-1/2}(x) \text{ is proportional to } \frac{\cosh x}{\sqrt{x}}.$$

5. Use asymptotic result (106) for $I_\nu(x)$ when x is large to find the constants of proportionality in Exercise 4, and then use the result of Exercise 2 to find $I_{3/2}(x)$ and $I_{-3/2}(x)$.
6. Use Lemma 6.1 of Section 6.3 to reduce Bessel's modified equation of order $\nu = 1/2$ to standard form, and hence show that when x is large two linearly independent solutions of the equation are proportional to $e^x/\sqrt{x}$ and $e^{-x}/\sqrt{x}$.
7. Deduce the expressions for $I_{\pm 1/2}(x)$ and $I_{\pm 3/2}(x)$ from the corresponding results for $J_{\pm 1/2}(x)$ and $J_{\pm 3/2}(x)$ in (72) to (74) of Section 8.6.
8. Use Abel's formula in Exercise 6 of set 6.1 to show that if y_1 and y_2 are any two linearly independent solutions of Bessel's modified equation, then

$$y_1 y_2' - y_2 y_1' = C/x,$$

where C is a constant introduced through the Abel formula.

9. Set $y_1(x) = I_\nu(x)$ and $y_2(x) = I_{-\nu}(x)$ in the result of Exercise 8, where ν is not an integer. Substitute the series for $I_\nu(x)$ and $I_{-\nu}(x)$, and by finding the coefficient of $1/x$ on the left-hand side identify the coefficient C. Use the result

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

to show that

$$I_\nu(x)I_{-\nu}'(x) - I_\nu'(x)I_{-\nu}(x) = -\frac{2}{\pi x}\sin \nu x.$$

10. Use the definition of $K_\nu(x)$ with the result of Exercise 9 to show that

$$I_\nu(x)K_\nu'(x) - I_\nu'(x)K_\nu(x) = -\frac{1}{x}.$$

11.* The amplitude $R(r)$ of the small symmetric vibrations of a flexible annular disc $a \le r \le b$ normal to its surface with its outer edge free and its inner edge fixed to a rod that oscillates along its length is governed by the equation

$$\frac{d^4R}{dr^4} + \frac{2}{r}\frac{d^3R}{dr^3} - \frac{1}{r^2}\frac{d^2R}{dr^2} + \frac{1}{r^3}\frac{dR}{dr} - R = 0.$$

Show by expressing the equation as

$$\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - 1\right)\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} + 1\right)R = 0$$

that its general solution is

$$R(r) = AJ_0(r) + BY_0(r) + CI_0(r) + DK_0(r),$$

where A, B, C, and D are arbitrary constants.

12.* In partial differential equations that govern physical phenomena with cylindrical and spherical polar coordinates, the following equation describes the radial variation $R(r)$ of the solution as a function of the radius r (see Chapter 18):

$$\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr} + \left(\lambda^2 - \frac{n^2}{r^2}\right)R = 0.$$

Here, λ is a parameter and $n = 0, 1, 2, \ldots$. Show that the general solution of the equation is

$$R(r) = AJ_n(\lambda r) + BY_n(\lambda r).$$

Find the form of the solution of the following boundary value problems, given that $R(r)$ remains bounded, and determine the permissible values of the parameter λ.

(i) $0 \le r \le a$, for all n with the boundary conditions $R(a) = 0$.
(ii) $b \le r \le c$, for all n with the boundary conditions $R(b) = R(c) = 0$.
(iii) $0 \le r \le a$, for all n with the boundary conditions $R(a) + kR'(a) = 0 (k = \text{const})$.
(iv) $b \le r \le c$, for $n = 0$ with the boundary conditions $R(b) = R'(c) = 0$.

8.9 A Critical Bending Problem: Is There a Tallest Flagpole?

The implication of the question posed in the section heading will have been experienced by anyone who has tried holding a long, thin, flexible rod in a vertical position. If the rod is short, and its tip is given a small sideways displacement and released, the rod will perform transverse oscillations until it reaches an equilibrium position in a bent shape because of supporting its own weight. The longer the rod, the larger the amplitude of these oscillations, and the greater the bending under its own weight when in equilibrium, until at some critical length the rod will bend until its tip just touches the ground, after which it will remain in that position.

Bessel functions and the bending of a thin vertical rod

An idealization of this phenomenon can be modeled by a long, thin, flexible flagpole of uniform cross-section, the base of which is clamped in the ground so the pole is vertical. We then ask at what length will the pole become unstable, so that any displacement of the top of the pole will cause it to bend under its own weight until the top of the pole touches and remains in contact with the ground? This question can be posed in mathematical terms, and it is the one that will be answered here.

The solution to this question will involve the use of Bessel functions, but the linear differential equation involved will have to satisfy a two-point boundary condition instead of the initial conditions we have considered so far. This means that the existence and uniqueness of solutions to initial value problems guaranteed by Theorem 6.2 no longer applies, so even when a solution can be found it may not be unique — more will be said about this later.

Let us model the problem by considering a thin uniform flexible rod of length L with a constant cross-section that is constructed from material with a Young's modulus of elasticity E, with the moment of inertia of a cross-section about a diameter normal to the plane of bending equal to I. The line density along the rod will be assumed to be constant and equal to w. The x-axis will be taken to be vertical and to coincide with the undistorted axis of the rod, with its origin located at the base of the rod. The horizontal displacement of the rod at a position x will be taken to be y, as shown in Fig. 8.10.

It is known from Section 5.2(f) that if the moment acting on the rod at a position x is $M(x)$, the equation governing its transverse deflection y when in equilibrium is

$$EI\frac{d^2y}{dx^2} = M(x). \tag{107}$$

The **shear** on the rod at point x is the force exerted perpendicular to the axis of the rod at x due to the weight of the rod extending from x to the top at P. As the length of this part of the rod is $L - x$, and its line density is w, the weight of this section is

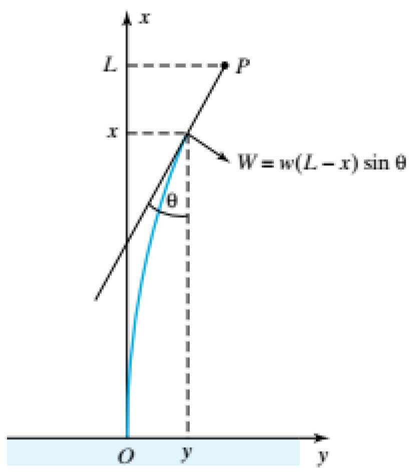


FIGURE 8.10 Equilibrium position of the rod when bent under its own weight.

given by $w(L-x)$, so the component W of this force normal to the axis of the rod at x is simply

$$W = w(L-x)\sin\theta, \tag{108}$$

where θ is the angle of deflection of the rod from the vertical at point x, as shown in Fig. 8.10.

It is known from mechanics that the shear on a rod is given in terms of the moment $M(x)$ by

$$\frac{dM}{dx} = -W(x). \tag{109}$$

We now make the approximation that the deflection at point x on the rod is small, so $\sin\theta \approx \tan\theta = dy/dx$, and by combining (107) to (109) we arrive that the governing equation for the deflection, which is the third order linear variable coefficient differential equation

$$EI\frac{d^3y}{dx^3} + w(L-x)\frac{dy}{dx} = 0. \tag{110}$$

Making the change of variable $z = L - x$ brings (110) to the more convenient form

$$\frac{d^3y}{dz^3} + \left(\frac{w}{EI}\right)z\frac{dy}{dz} = 0. \tag{111}$$

To apply this to our problem it is necessary to determine appropriate boundary conditions to be applied at the base and top of the rod. An obvious condition to be applied at the base is that due to clamping the pole in a vertical position at the origin, $(dy/dx)_{x=0} = (dy/dz)_{z=L} = 0$. To arrive at a second condition we notice that when the rod is bent and in equilibrium, there can be no bending moment at the top of the rod, so it can have no curvature at that point. Recalling that the radius of curvature ρ of a plane curve $y = y(x)$ is

$$\rho = \frac{(1+(y')^2)^{2/3}}{y''}, \tag{112}$$

we see that the rod will have no curvature at $x = L$ (equivalently at $z = 0$) when $\rho = \infty$, corresponding to $(d^2y/dx^2)_{x=L} = (d^2y/dz^2)_{z=0} = 0$.

Setting $u(z) = dy/dz$, these two boundary conditions become

$$u(L) = 0 \quad \text{and} \quad (du/dz)_{z=0} = 0. \tag{113}$$

Equation (111) is third order, but in terms of $u(z)$ it is only second order, and we have found two conditions on $u(z)$ from which to determine u. Fortunately, we only need to work with $u(z)$ to solve our problem. This is because we will soon see that the two-point boundary conditions (113) applied to the differential equation for u

$$\frac{d^2u}{dz^2} + \left(\frac{w}{EI}\right)zu = 0 \tag{114}$$

will provide sufficient information for us to find the critical length at which bending occurs.

Identifying equation (114) with (93) from Section 8.7, with x replaced by z, shows that

$$1 - 2a = 0, \quad 2c - 2 = 1, \quad a^2 - \nu^2c^2 = 0, \quad \text{and} \quad b^2c^2 = w/EI, \tag{115}$$

so

$$a = 1/2, \quad c = 3/2, \quad \nu = 1/3, \quad \text{and} \quad b = \frac{2}{3}\sqrt{\frac{w}{EI}}. \tag{116}$$

Using this information in the solution (94) to equation (93) in Section 8.7 gives

$$u(z) = C_1\sqrt{z}\,J_{1/3}\left(\frac{2}{3}\sqrt{\frac{w}{EI}}z^{3/2}\right) + C_2\sqrt{z}\,J_{-1/3}\left(\frac{2}{3}\sqrt{\frac{w}{EI}}z^{3/2}\right). \tag{117}$$

Noticing from (71) of Section 8.6 that for small z

$$J_\nu(z) \approx \frac{1}{\Gamma(1+\nu)}\left(\frac{z}{2}\right)^{\nu} \quad \text{and} \quad J_{-\nu}(z) \approx \frac{1}{\Gamma(1-\nu)}\left(\frac{z}{2}\right)^{-\nu},$$

we see that close to the top of the rod, that is, for small z, $u(z)$ can be approximated by

$$u(z) \approx C_1\frac{z}{\Gamma(4/3)}\left(\frac{1}{3}\sqrt{\frac{w}{EI}}\right)^{1/3} + C_2\frac{1}{\Gamma(2/3)}\left(\frac{1}{3}\sqrt{\frac{w}{EI}}\right)^{-1/3}.$$

Differentiation of this result gives

$$u'(z) \approx C_1\frac{1}{\Gamma(4/3)}\left(\frac{1}{3}\sqrt{\frac{w}{EI}}\right)^{1/3},$$

but to satisfy the second boundary condition $(du/dz)_{z=0} = 0$, we must set $C_1 = 0$, causing solution (117) to reduce to

$$u(z) = C_2\sqrt{z}\,J_{-1/3}\left(\frac{2}{3}\sqrt{\frac{w}{EI}}z^{3/2}\right). \tag{118}$$

Applying the remaining boundary condition $u(L) = 0$ to (118) gives

$$0 = C_2\sqrt{L}\,J_{-1/3}\left(\frac{2}{3}\sqrt{\frac{w}{EI}}L^{3/2}\right), \tag{119}$$

and this will be satisfied if either $C_2 = 0$ or $J_{-1/3}(\frac{2}{3}\sqrt{\frac{w}{EI}}L^{3/2}) = 0$. The first condition $C_2 = 0$ corresponds to the unstable equilibrium configuration in which the rod is vertical, and so must be rejected, whereas the second condition corresponds to the required critical bending condition, and it will be satisfied when L is such that it causes $J_{-1/3}$ to vanish.

It is at this stage that we discover the boundary value problem does *not* have a unique solution, because the asymptotic behavior of $J_{-1/3}$ given in (70) of Section 8.6 shows that it has infinitely many zeros. To resolve this difficulty, and to find the length at which critical bending occurs, we must now seek a selection criterion for the length from *outside* the description of the physical situation provided by the differential equation.

Such a criterion is not hard to find, because critical bending must occur at the *smallest* value of L, say at L_c, that satisfies the condition

$$J_{-1/3}\left(\frac{2}{3}\sqrt{\frac{w}{EI}}L_c^{3/2}\right) = 0, \tag{120}$$

because if critical bending occurs when $L = L_c$, it will certainly occur at any larger value of L.

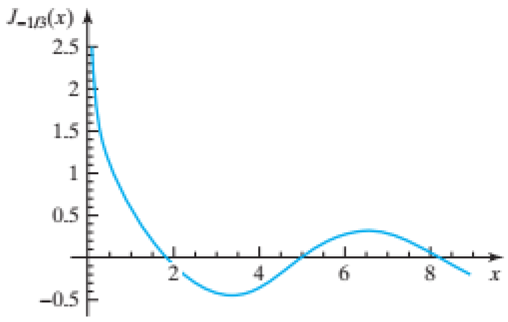


FIGURE 8.11 Graph of $J_{-1/3}(x)$ showing its first few zeros.

A graph of $J_{-1/3}(x)$ is shown in Fig. 8.11, from which it can be seen that the first zero α of $J_{-1/3}(x)$ occurs at around the value $\alpha \approx 1.87$, though numerical calculation provides the more accurate value $\alpha = 1.86635\ldots$. However, this accuracy is unnecessary, because the approximations made when modeling the physical situation introduce errors of sufficient magnitude that the value $\alpha \approx 1.87$ is adequate.

Using the value $\alpha = 1.87$ shows that the length L_c for critical bending must satisfy the formula

$$\frac{2}{3}\sqrt{\frac{w}{EI}}L_c^{3/2} \approx 1.87,$$

which is equivalent to

$$L_c \approx 1.99\left(\frac{EI}{w}\right)^{1/3}.$$

This approximation shows, as would be expected, that if the rod is not cylindrically symmetric about its axis, the critical length L_c will depend on the plane in which bending occurs, because the moment of inertia will depend on the direction in which the rod bends. Thus, for example, the critical length of a rod with a rectangular cross-section that bends in a plane parallel to one pair of its faces will differ from the critical length when bending occurs in a plane parallel to its other pair of faces. In such cases the model used is too simple, because twisting (torsion) will be likely to occur, causing the rod always to buckle in such a way that L_c assumes its smallest possible value.

The simplest case arises when the rod has a circular cross-section of radius a, for then the moment of inertia of the cross-section about any diameter is $I = \pi a^4/4$. When this expression is substituted into the approximation for L_c, we obtain the approximation

$$L_c \approx 1.25\left(\frac{Ea^4}{w}\right)^{1/3}.$$

Summary

In addition to involving Bessel functions, this idealization of a physical problem has illustrated the way in which a mathematical approach can sometimes lead to more than one solution, only one of which can be regarded as an approximation to the situation in the real world. The choice of the appropriate solution was seen to be based on an additional physical consideration that was outside the original formulation of the mathematical problem. This situation is not unusual in applied mathematics, where the choice of solution is often based on stability considerations, a physically possible solution being stable, whereas a nonphysical solution is unstable and so will not be observed. A different example occurs in the study of shock waves in air where two solutions are mathematically possible, though only one is physically realizable. In that case the selection principle is based on the thermodynamics of the problem, though it can also be based on stability considerations.

8.10 Sturm–Liouville Problems, Eigenfunctions, and Orthogonality

Mathematical models of physical situations arising in engineering and physics lead to two-point boundary value problems for a function $y(x)$ that is defined over an interval $a < x < b$ and satisfies a differential equation of the form

$$y''(x) + P(x)y'(x) + (Q(x) + \lambda R(x))y(x) = 0, \tag{121}$$

in which λ is a parameter. This equation always has the solution $y(x) \equiv 0$, called the **trivial solution**, but if it is to have nontrivial solutions (solutions that are not identically zero) satisfying boundary conditions at $x = a$ and $x = b$, the parameter λ cannot be arbitrary. In what follows our purpose will be to find constant values of λ for which nontrivial solutions exist satisfying given boundary conditions. It will be seen later how these nontrivial solutions can be used to generalize series expansions of arbitrary functions over the interval $a < x < b$ that, along with other uses, are needed in Chapter 18 when solving partial differential equations by the method of separation of variables.

To proceed further we will write (121) in a more convenient form, and to this end we simplify its first two terms using the method developed in Section 5.6 when finding an integrating factor for a linear first order equation. Defining the function $p(x)$ as

$$p(x) = \exp\left[\int P(x)\mathrm{d}x\right],$$

and multiplying (121) by $p(x)$ gives

$$p(x)[y''(x) + P(x)y'(x)] + p(x)(Q(x) + \lambda R(x))y(x) = 0.$$

However,

$$p(x)[y''(x) + P(x)y'(x)] = \frac{d}{dx}\left[p(x)\frac{dy(x)}{dx}\right],$$

so the equation becomes

$$\frac{d}{dx}\left[p(x)\frac{dy(x)}{dx}\right] + p(x)(Q(x) + \lambda R(x))y(x) = 0.$$

Finally, setting $q(x) = p(x)Q(x)$ and $r(x) = p(x)R(x)$ allows equation (121) to be written in the form

$$\frac{d}{dx}\left[p(x)\frac{dy(x)}{dx}\right] + [q(x) + \lambda r(x)]y(x) = 0. \tag{122}$$

In what follows $p(x), q(x), r(x)$, and $p'(x)$ will be assumed to be continuous functions defined on a closed interval $a \le x \le b$ on which $p(x) > 0, r(x) > 0$.

Differential equations with these properties and written in this form are called **Sturm–Liouville equations**, and the type of boundary conditions that are to be imposed will be introduced after the following typical examples of these equations.

JACQUES CHARLES FRANÇOIS STURM (1803–1855) AND JOSEPH LIOUVILLE (1809–1882)
Sturm, who was born in Geneva, Switzerland, was Poisson's successor in the Chair of Mechanics in the Sorbonne. Much of his work was in algebra, where he worked on the determination of intervals on the real line inside each of which was located one real root of a polynomial, though he also worked on the study of heat flow introduced by his contemporary Joseph Fourier. Liouville, a professor at the Collège de France, also studied algebraic problems and, in particular, quadratic forms, though he also made contributions to elliptic functions and to complex analysis. Sturm and Liouville, who were friends, collaborated on the eigenvalue and eigenfunction problems raised by the study of heat flow, and together their work led to what is now called the study of Sturm–Liouville systems.

examples of Sturm–Liouville equations

Simple harmonic motion equation

The differential equation describing undamped simple harmonic oscillations

$$y'' + n^2 y = 0 \tag{123}$$

follows from (122) by setting $p(x) = 1, q(x) = 0, r(x) = 1$, and $\lambda = n^2$.

The Legendre equation

The Legendre equation encountered in (10) of Section 8.2, usually written

$$(1 - x^2)y'' - 2xy' + \alpha(\alpha + 1)y = 0, \tag{124}$$

follows from (122) by setting $p(x) = 1 - x^2, q(x) = 0, r(x) = 1$, and $\lambda = \alpha(\alpha + 1)$.

Bessel's equation

When Bessel's equation of order ν is written in its more general form

$$x^2y'' + xy' + (k^2x^2 - \nu^2)y = 0, \tag{125}$$

the equation follows from (122) by setting $p(x) = x, q(x) = -\nu^2/x, r(x) = x$, and $\lambda = k^2$.

The Chebyshev equation

The Chebyshev equation of order ν is

$$(1 - x^2)y'' - xy' + n^2y = 0, \tag{126}$$

and the equation follows from (122) by setting $p(x) = (1 - x^2)^{1/2}, q(x) = 0, r(x) = (1 - x^2)^{-1/2}$, and $\lambda = n^2$.

For future reference, Table 8.3 lists $p(x), q(x), r(x)$, and λ for the preceding equations, together with three other named equations that find applications in numerical analysis and elsewhere.

TABLE 8.3 $p(x)$, $q(x)$, $r(x)$ and λ for Some Named Equations

Name	$p(x)$	$q(x)$	$r(x)$	λ
Simple harmonic equation	1	0	1	n^2
Legendre's equation	$1-x^2$	0	1	$\alpha(\alpha+1)$
Bessel's equation	x	$-\nu^2/x$	x	k^2
Bessel's modified equation	x	$-\nu^2/x$	$-x$	k^2
Laguerre equation	xe^{-x}	0	e^{-x}	n
Chebyshev equation	$(1-x^2)^{1/2}$	0	$(1-x^2)^{-1/2}$	n^2
Hermite equation	e^{-x^2}	0	e^{-x^2}	$2n$

When the Sturm–Liouville equation (122) is associated with boundary conditions at $x = a$ and $x = b$, the equation itself together with the boundary conditions form what is called a **Sturm–Liouville problem**. The boundary conditions that will concern us here are the **homogeneous** boundary conditions,

$$A_1 y(a) + A_2 y'(a) = 0 \quad \text{and} \quad B_1 y(b) + B_2 y'(b) = 0, \tag{127}$$

where the term *homogeneous* is used in the sense that the linear combinations of $y(x)$ and $y'(x)$ at $x = a$ and $x = b$ are both equal to zero. There are three categories of Sturm–Liouville problems that arise, called **regular**, **periodic**, and **singular** problems according to the nature of the boundary conditions and the behavior of $p(x)$ at the boundaries.

Regular Sturm–Liouville problems

Regular problems are those for which constant values of λ are sought corresponding to each of which a nontrivial solution can be found for the Sturm–Liouville equation

$$(py')' + (q + \lambda r)y = 0,$$

with $p(x) > 0$ continuous on $a \le x \le b$ and subject to the boundary conditions

$$A_1 y(a) + A_2 y'(a) = 0 \quad \text{and} \quad B_1 y(b) + B_2 y'(b) = 0,$$

where in neither of the boundary conditions do both constant coefficients vanish.

Periodic Sturm–Liouville problems

This class of problems arises when $p(x)$ and the boundary conditions involving $y(x)$ and $y'(x)$ are periodic over the interval $a \le x \le b$. In this case constant values of λ are sought corresponding to each of which a nontrivial solution can be found for the Sturm–Liouville problem

$$(py')' + (q + \lambda r)y = 0,$$

subject to the periodic boundary conditions

$$p(a) = p(b), \quad y(a) = y(b), \quad \text{and} \quad y'(a) = y'(b).$$

Singular Sturm–Liouville problems

In this class of problems constant values of λ are sought, corresponding to each of which a nontrivial solution can be found for the Sturm–Liouville equation

$$(py')' + (q + \lambda r)y = 0,$$

on a finite interval at one or or both ends of which $p(x)$ or $r(x)$ vanish, or on a semi-infinite or infinite interval. The most frequently occurring problem of this type, and the only one to be considered here, is the Sturm–Liouville problem defined on a finite interval $a \le x \le b$, where the singular point is located at either $x = a$ or $x = b$, so that either $p(a) = 0$ or $p(b) = 0$. In such cases the boundary condition that is often imposed at the singular point takes the form of the requirement that the solution remains bounded there. Typically, this happens when a bounded solution of Bessel's equation of the form $y(x) = AJ_0(x) + BY_0(x)$ is required over an interval $0 \le x \le a$, because then the requirement that the solution remains bounded at the singular point located at $x = 0$ means we must set $B = 0$ to exclude the infinite value of $Y_0(x)$ at $x = 0$.

When dealing with Sturm–Liouville problems, each value of λ for which a nontrivial solution can be found is called an **eigenvalue** of the problem, and the corresponding solution $y(x)$ is called an **eigenfunction** of the problem. Because the Sturm–Liouville equation (122) is homogeneous, it follows that an eigenfunction can be multiplied by any constant factor and still remain an eigenfunction. This simple but fundamental property will be used repeatedly, first when normalizing eigenfunctions and later when representing arbitrary functions defined over an interval $[a, b]$ in terms of series of eigenfunctions, as is done in Chapter 9 when working with Fourier series. Such representations of functions are called *eigenfunction expansions*.

In most practical situations an eigenvalue is associated with an important physical characteristic of the problem, such as the frequency of vibration of a string or of a metal plate. In such cases the eigenfunction can be considered to describe a "snapshot" of a particular mode of vibration of the string or plate when it vibrates at the frequency determined by the associated eigenvalue. This application, and others that lead to Sturm–Liouville problems, will be developed in detail when partial differential equations are discussed in the context of *separation of variables*.

A Regular Problem

EXAMPLE 8.18 Find the eigenvalues and eigenfunctions of the two-point boundary value problem

$$y'' + \lambda y = 0,$$

such that

$$y(0) = 0 \quad \text{and} \quad y'(\pi) = 0.$$

Solution The interval over which the eigenfunctions are defined is $0 \le x \le \pi$. We need to consider the three cases $\lambda = 0, \lambda < 0$, and $\lambda > 0$. The homogenous boundary conditions in this problem are of the type given in (127) with $A_2 = 0$ and $B_1 = 0$, where the values of the constants A_1 and B_2 are immaterial provided neither is zero.

***Case* $\lambda = 0$**

When $\lambda = 0$ the equation has the general solution

$$y(x) = C_1 x + C_2,$$

so to satisfy the boundary condition $y(0) = 0$ we must have $C_2 = 0$, and to satisfy the boundary condition $y'(\pi) = 0$ we must have $C_1 = 0$, giving rise to the trivial solution $y(x) \equiv 0$. Thus, $\lambda = 0$ is not an eigenvalue of the problem.

***Case* $\lambda < 0$**

If we set $\lambda = -\mu^2$, the general solution becomes

$$y(x) = C_1 e^{\mu x} + C_2 e^{-\mu x},$$

so the imposition of the boundary conditions requires that

$$0 = C_1 + C_2 \quad \text{and} \quad 0 = \mu C_1 e^{\mu\pi} - \mu C_2 e^{-\mu\pi}.$$

After the elimination of C_2, this last result can be written

$$0 = 2\mu C_1 \cosh \mu\pi,$$

but $\mu > 0$, so as $\cosh \mu\pi \neq 0$, this is only possible if $C_1 = 0$, so again we obtain the trivial solution showing that the problem has no negative eigenvalues.

***Case* $\lambda > 0$**

As $\lambda > 0$, it is convenient to set $\lambda = \mu^2$, when the general solution of the equation becomes

$$y(x) = C_1 \cos \mu x + \mu C_2 \sin \mu x.$$

Applying the boundary condition $y(0) = 0$ to the general solution gives $C_1 = 0$, and applying the boundary condition $y'(\pi) = 0$ gives

$$\mu C_2 \cos \mu\pi = 0,$$

so either $C_2 = 0$ or $\cos \mu\pi = 0$. If we take $C_2 = 0$, then as $C_1 = 0$ we obtain the trivial solution, so we must take $C_2 \neq 0$. The condition $\cos \mu\pi = 0$ is satisfied if $\mu\pi$ is one of the zeros of the cosine function given by $\pm\frac{1}{2}(2n+1)\pi$, for $n = 0, 1, 2, \ldots$.

Denoting the permitted values of μ by μ_n we arrive at the condition

$$\mu_n = \pm\frac{1}{2}(2n+1), \quad \text{with } n = 0, 1, 2, \ldots.$$

The eigenvalues of this problem corresponding to the parameter $\lambda = \mu^2$ are thus

$$\lambda_n = \frac{(2n+1)^2}{4}, \quad \text{with } n = 0, 1, 2, \ldots,$$

and the corresponding eigenfunctions are

$$y_n(x) = \sin \frac{(2n+1)x}{2} \quad \text{with } n = 0, 1, 2, \ldots.$$

When writing down the form of the eigenfunction $y_n(x)$, we have set $C_2 = 1$ because, as has already been remarked, an eigenfunction can be multiplied by any constant nonzero factor and still remain an eigenfunction.

This example has shown the existence of an infinite increasing sequence of positive eigenvalues μ_n^2, corresponding to each of which there is a nontrivial solution of the Sturm–Liouville problem, namely the eigenfunction $y_n(x) = \sin \mu_n x$. If $\mu \neq \mu_n$, then the Sturm–Liouville problem only has the *trivial* solution $y(x) \equiv 0$. ■

A Periodic Problem

EXAMPLE 8.19 Find the eigenvalues and eigenfunctions of the Sturm–Liouville equation

$$y'' + \lambda y = 0$$

subject to the conditions

$$y(0) = y(L), \quad y'(0) = y'(L).$$

Solution The interval over which the eigenfunctions are defined is $0 \leq x \leq L$, and as in Example 8.18 we must again consider the three cases $\lambda = 0$, $\lambda < 0$, and $\lambda > 0$.

Case $\boldsymbol{\lambda = 0}$

As in the previous problem, the general solution is

$$y(x) = C_1 x + C_2,$$

so applying the boundary condition $y(0) = y(L)$ leads to the result $C_2 = C_1 L + C_2$, from which it follows that $C_1 = 0$. As $y'(x) = C_1$ the boundary condition $y'(0) = y'(L)$ is automatically satisfied, showing that $y(x) = C_2$, with C_2 any nonzero constant. This shows that in this case $\lambda = 0$ is an eigenvalue, and that $y(x) = C_2$ (C_2 is an arbitrary nonzero constant) is the corresponding eigenfunction.

Case $\boldsymbol{\lambda < 0}$

If we set $\lambda = -\mu^2$, the general solution becomes

$$y(x) = C_1 e^{\mu x} + C_2 e^{-\mu x}.$$

The boundary condition $y(0) = y(L)$ leads to the condition

$$C_1(1 - e^{\mu L}) = C_2(e^{-\mu L} - 1),$$

and the boundary condition $y'(0) = y'(L)$ leads to the condition

$$C_1(1 - e^{\mu L}) = -C_2(e^{-\mu L} - 1).$$

This last condition is only possible if $C_1 = 0$, but then $C_2 = 0$, so we again obtain the trivial solution. Consequently, we conclude that this problem has no negative eigenvalues.

Case $\boldsymbol{\lambda > 0}$

Setting $\lambda = \mu^2$ the general solution of the equation becomes

$$y(x) = C_1 \cos \mu x + C_2 \sin \mu x.$$

The boundary condition $y(0) = y(L)$ leads to the condition

$$C_1(1 - \cos \mu L) = C_2 \sin \mu L,$$

and the boundary condition $y'(0) = y'(L)$ leads to the condition

$$C_2(1 - \cos \mu L) = -C_1 \sin \mu L.$$

Eliminating C_2 between these two equations and simplifying the result gives

$$2C_1(1 - \cos \mu L) = 0.$$

This condition is satisfied if either $C_1 = 0$, or if $\cos \mu L = 1$. If $C_1 = 0$, then $C_2 = 0$, and we obtain the trivial solution, so the only other possibility is that $\cos \mu L = 1$. This last condition will be satisfied if μL is zero or an integer multiple of 2π, so

$$\mu L = \pm 2n\pi \quad \text{for } n = 0, 1, 2, \ldots,$$

or

$$\mu_n = \pm 2n\pi/L \quad \text{for } n = 0, 1, 2, \ldots.$$

As $\lambda = \mu^2$ the eigenvalues are seen to be

$$\lambda_n = 4n^2\pi^2/L^2, \quad \text{for } n = 0, 1, 2, \ldots.$$

The corresponding eigenfunctions are

$$y_n(x) = C_1 \cos \mu_n x + C_2 \sin \mu_n x,$$

or

$$y_n(x) = C_1 \cos(2n\pi x/L) + C_2 \sin(2n\pi x/L), \quad \text{for } n = 0, 1, 2, \ldots,$$

where not both constants C_1 and C_2 are zero. Because C_1 and C_2 are arbitrary, and both the cosine function and the sine function satisfy the Sturm–Liouville equation and the boundary conditions, by first setting $C_1 = 1$ and $C_2 = 0$ and then $C_1 = 0$ and $C_2 = 1$ it is seen that in this case the *single* eigenvalue $\lambda_n = 4n^2\pi^2/L^2$ has associated with it the *two* distinct eigenfunctions

$$y_n^{(1)}(x) = \cos(2n\pi x/L) \quad \text{and} \quad y_n^{(2)}(x) = \sin(2n\pi x/L).$$ ■

The eigenvalues in Sturm–Liouville problems are not always determined as easily as in the previous examples, and this is illustrated by the next example.

EXAMPLE 8.20 Find the eigenvalues and eigenfunctions of the Sturm–Liouville equation

$$y'' + \lambda y = 0,$$

subject to the conditions

$$y(0) - y'(0) = 0, \quad y(1) + y'(1) = 0.$$

Solution The interval over which the eigenfunctions are defined is $0 \le x \le 1$, and as before we must again consider the cases $\lambda = 0$, $\lambda < 0$, and $\lambda > 0$.

Case $\lambda = 0$

The general solution is

$$y(x) = C_1 x + C_2,$$

so applying the boundary condition $y(0) - y'(0) = 0$ shows that $C_2 - C_1 = 0$, while applying the boundary condition $y(1) + y'(1) = 0$ gives the condition $2C_1 + C_2 = 0$.

The only solution for these equations is $C_1 = C_2 = 0$ corresponding to the trivial solution, so $\lambda = 0$ is not an eigenvalue of the problem.

***Case* $\lambda < 0$**

Setting $\lambda = -\mu^2$ leads to the general solution

$$y(x) = C_1 e^{\mu x} + C_2 e^{-\mu x}.$$

Applying the boundary condition $y(0) - y'(0) = 0$ leads to the condition

$$C_1(1-\mu) + C_2(1+\mu) = 0,$$

and applying the boundary condition $y(1) + y'(1) = 0$ leads to the condition

$$C_1[(1+\mu)e^{\mu} + C_2(1-\mu)e^{-\mu}] = 0.$$

As a factor $\mu - 1$ appears, we must consider the cases $\mu = 1$ and $\mu \neq 1$ separately. If $\mu = 1$, the first equation gives $C_2 = 0$, and the second one gives $C_1 = 0$, corresponding to the trivial solution. So $\mu = 1$ is not an eigenvalue. If $\mu \neq 1$, eliminating C_2 between these two equations leads to the condition

$$C_1[(1+\mu)^2 e^{\mu} - (1-\mu)^2 e^{-\mu}] = 0.$$

As $\mu > 0$, $(\mu+1)^2 e^{\mu} > (\mu-1)^2 e^{-\mu}$, showing that the bracketed term is nonvanishing, from which we conclude that $C_1 = 0$, and so $C_2 = 0$, corresponding to the trivial solution. Thus, this Sturm–Liouville problem has no negative eigenvalues.

***Case* $\lambda > 0$**

Setting $\lambda = \mu^2$ leads to the general solution

$$y(x) = C_1 \cos \mu x + C_2 \sin \mu x.$$

Applying the boundary condition $y(0) - y'(L) = 0$ shows that

$$C_1 - \mu C_2 = 0,$$

and applying the boundary condition $y(1) + y'(1) = 0$ gives

$$C_1 \cos \mu + C_2 \sin \mu - \mu C_1 \sin \mu + \mu C_2 \cos \mu = 0.$$

Eliminating C_1 between these two equations, we obtain

$$C_2[2\mu \cos \mu + (1-\mu^2)\sin \mu] = 0.$$

The constant C_2 cannot be zero, because then $C_1 = 0$, corresponding to the trivial solution, so μ must be a solution of the equation

$$2\mu \cos \mu + (1-\mu^2)\sin \mu = 0$$

or, equivalently, μ_n is a solution of the transcendental equation

$$\tan \mu_n = \frac{2\mu_n}{\mu_n^2 - 1}.$$

This equation can only be solved numerically, but approximate solutions can be found graphically. Figure 8.12(a) shows graphs of $y = \tan \mu$ and $y = 2\mu/(\mu^2 - 1)$, and the required solutions μ_n are the values of μ at which the graphs intersect. It has been shown that $\mu = 1$ is not an eigenvalue, so the permissible values of μ_n are all greater than 1. The vertical lines to the right of $x = 1$ are the asymptotes to the

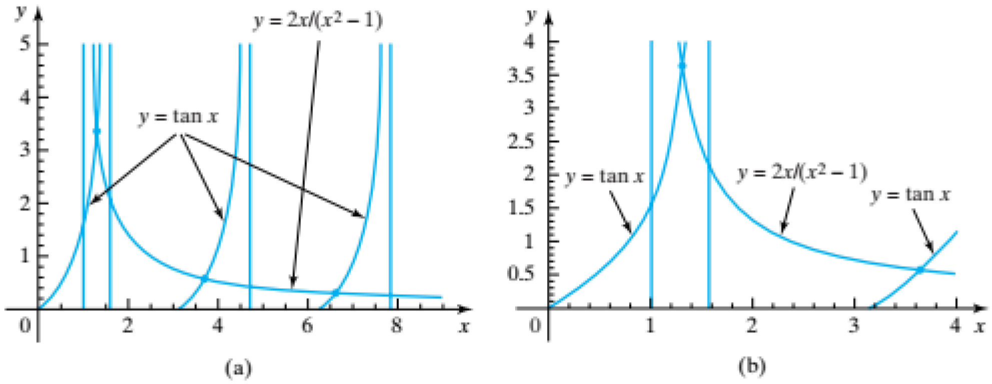


FIGURE 8.12 The roots of $\tan \mu = 2\mu/(\mu^2 - 1)$.

tangent function, and the vertical line at $x = 1$ is the asymptote to $2x/(x^2 - 1)$, to the right of which must lie all the solutions μ_n. The graph in Fig. 8.12b, drawn on a larger scale, shows that the first two values of μ are approximately $\mu_1 = 1.3$ and $\mu_2 = 3.7$. A numerical calculation using Newton's method described in Chapter 19 gives the better approximations $\mu_1 = 1.30654$ and $\mu = 3.67319$. It can be seen from Fig. 8.12a that when n is large $\mu_n \approx n\pi$. ■

A Singular Problem

EXAMPLE 8.21 Find the eigenvalues and eigenfunctions of Bessel's equation

$$x^2 y'' + xy' + (k^2x^2 - n^2)y = 0$$

on the interval $0 \le x \le a$ on which the solution is bounded with $y(a) = 0$.

Solution This is a singular Sturm–Liouville problem, because when Bessel's equation is written in the Sturm–Liouville form

$$\frac{d}{dx}\left[x\frac{dy}{dx}\right] + \left(k^2x^2 - \frac{n^2}{x}\right)y = 0,$$

with $p(x) = x$, $q(x) = -n^2/x$, $r(x) = x$, and $\lambda = k^2$ (see Table 8.3), it is seen that $p(0) = 0$.

The general solution is

$$y(x) = C_1 J_n(kx) + C_2 Y_n(kx),$$

but $Y_n(kx)$ is infinite when $x = 0$, so for the solution to remain finite over the interval $0 \le x \le a$ we must set $C_2 = 0$.

The solution now reduces to

$$y(x) = C_1 J_n(kx),$$

so if the boundary condition $y(a) = 0$ is to be satisfied we must set

$$J_n(ka) = 0.$$

This condition will be satisfied if ka is one of the zeros of $J_n(x)$. If we denote the zeros of $J_n(x)$ by $j_{n,r}$, with $r = 1, 2, \ldots$, it follows that k must be such that it assumes one of the values

$$k_n = j_{n,r}/a, \quad \text{with } r = 1, 2, \ldots.$$

Thus, the eigenvalues $\lambda_n = k_n^2$ are given by

$$\lambda_n = j_{n,r}^2/a^2,$$

and the corresponding eigenfunctions are

$$y_r(x) = J_n(j_{n,r}x/a), \quad \text{with } r = 1, 2, \ldots,$$

where for convenience we have set $C_1 = 1$. Table 8.1 lists the first six zeros of $J_n(x)$ for $n = 0, 1, 2, 3$. Thus if, for example, we consider the case $n = 0$, the corresponding zeros are seen to be $j_{0,1} = 2.4048$, $j_{0,2} = 5.5201\ldots$, so the eigenvalues are $\lambda_1 = 5.7832/a^2$, $\lambda_2 = 30.4711/a^2, \ldots$, and the corresponding eigenfunctions are

$$y_1(x) = J_0(2.4048x/a), \quad y_2(x) = J_0(5.5201x/a), \ldots.$$

■

Orthogonal and Orthonormal Systems of Functions

orthogonal and orthonormal systems

When working with eigenfunctions it is useful to introduce the notions of **orthogonal** and **orthonormal** systems of eigenfunctions that are defined as follows.

Let $\varphi_1(x), \varphi_2(x), \ldots$ be an infinite sequence of functions defined over the interval $a \le x \le b$ on which a function $r(x) \ge 0$ is defined. Then the functions are said to be **orthogonal** with respect to the **weight function** $r(x)$ if

$$\int_a^b r(x)\varphi_m(x)\varphi_n(x)dx = 0 \quad \text{for } m \ne n.$$

Clearly, the integral $\int_a^b r(x)\varphi_m(x)\varphi_n(x)dx > 0$ when $m = n$, so we can define a number $\|\varphi_n(x)\|$, called the **norm** of $\varphi_n(x)$, where the square of the norm is defined as

$$\|\varphi_n(x)\|^2 = \int_a^b r(x)\varphi_n^2(x)dx.$$

Using this definition of the norm it is easy to see that the sequence of normalized functions $\hat{\varphi}_1(x) = \varphi_1(x)/\|\varphi_1(x)\|$, $\hat{\varphi}_2(x) = \varphi_2(x)/\|\varphi_2(x)\|, \ldots$ has the property that

$$\int_a^b \hat{\varphi}_m(x)\hat{\varphi}_n(x)r(x)dx = 0, \quad \text{for } m \ne n$$

and

$$\int_a^b \hat{\varphi}_m(x)\hat{\varphi}_n(x)r(x)dx = 1, \quad \text{for } m = n.$$

The sequence of functions $\hat{\varphi}_1(x), \hat{\varphi}_2(x), \ldots$ derived from the sequence of orthogonal functions $\varphi_1(x), \varphi_2(x), \ldots$ by normalization is said to form an **orthonormal** sequence of functions.

In what follows the orthogonality of eigenfunctions will be used extensively, but for the moment it will be sufficient to give a single elementary example of an orthogonal sequence of functions.

EXAMPLE 8.22 Show that the sequence of functions

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cos 3x,\ \sin 3x, \ldots$$

is orthogonal over the interval $-\pi \leq x \leq \pi$ with respect to the weight function $r(x) = 1$, and use it to construct an orthonormal sequence.

Solution The functions in this sequence occur in the Fourier series representation of an arbitrary function $f(x)$ defined over the interval $-\pi \leq x \leq \pi$ that is discussed in Chapter 9. Routine calculation shows that for $m \neq n$,

$$\int_{-\pi}^{\pi} \sin mx \sin nx\, dx = 0, \quad \int_{-\pi}^{\pi} \cos mx \cos nx\, dx = 0, \quad \int_{-\pi}^{\pi} \sin mx \cos nx\, dx = 0,$$

and

$$\int_{-\pi}^{\pi} 1 dx = 2\pi, \quad \int_{-\pi}^{\pi} \sin^2 nx dx = \int_{-\pi}^{\pi} \cos^2 nx dx = \pi, \quad n = 1, 2, \ldots,$$

while $\int_{-\pi}^{\pi} 1 \cdot \cos mx dx = \int_{-\pi}^{\pi} 1 \cdot \sin mx dx = 0$.

So the functions

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cos 3x,\ \sin 3x, \ldots$$

are orthogonal over the interval $-\pi \leq x \leq \pi$ with respect to the weight function $r(x) = 1$. The respective norms are $\|1\| = \sqrt{2\pi}$ and $\|\sin nx\| = \|\cos nx\| = \sqrt{\pi}$, so the sequence of functions

$$1/\sqrt{2\pi}, \quad (\sin nx)/\sqrt{\pi}, \quad (\cos nx)/\sqrt{\pi}, \quad \text{with } n = 1, 2, \ldots,$$

forms an orthonormal sequence. ■

Fundamental Properties of Eigenvalues

The theorem that follows lists the most important properties of the eigenvalues and eigenfunctions of Sturm–Liouville problems. Apart from the important Rayleigh quotient that occurs in Theorem 8.3 (5), the other properties are all qualitative and their main use is to provide general information about eigenvalues that is often of considerable value when working with physical problems.

For convenience, the proofs of all results in Theorem 8.3 that can be established in a straightforward manner have been included in an appendix at the end of this chapter. The proofs of the other results can be found in the references listed at the end of the chapter. A reader who does not require the proofs that are given here may omit them, though the properties themselves should be understood.

THEOREM 8.3 **A Sturm–Liouville theorem**

important properties of eigenvalues

1. Regular and periodic Sturm–Liouville problems have an infinite number of distinct real eigenvalues $\lambda_1, \lambda_2, \ldots,$ that can be arranged in order so that

$$\lambda_1 < \lambda_2 < \lambda_3 < \ldots,$$

where the smallest eigenvalue λ_1 is finite, and

$$\lim_{n\to\infty} \lambda_n = \infty.$$

2. To each eigenvalue of a regular Sturm–Liouville problem there corresponds only one eigenfunction that is unique apart from an arbitrary multiplicative constant.
3. Let the eigenfunctions of a Sturm–Liouville problem on an interval $a \le x \le b$ with weight function $r(x)$ be denoted by $\varphi_1, \varphi_2, \ldots$, with the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots$. Then, if φ_m and φ_n are eigenfunctions corresponding to two distinct eigenvalues λ_m and λ_n ($\lambda_m \ne \lambda_n$ for $m \ne n$), the functions are orthogonal with respect to the weight function $r(x)$, so

$$\int_a^b r(x)\varphi_m(x)\varphi_n(x)dx = 0.$$

4. All the eigenvalues of a Sturm–Liouville problem are real.
5. Let λ_n be an eigenvalue of a regular Sturm–Liouville problem, with φ_n its associated eigenfunction defined on an interval $a \le x \le b$. Then λ_n is given in terms of the Sturm–Liouville functions p, q, r, and the boundary conditions by the **Rayleigh quotient**

$$\lambda_n = \frac{-[p\varphi_n\varphi_n']_a^b + \int_a^b p(\varphi_n')^2 dx - \int_a^b q\varphi_n^2\, dx}{\int_a^b r\varphi_n^2\, dx}.$$

6. Let λ_n be an eigenvalue and φ_n be the corresponding eigenfunction of a regular Sturm–Liouville problem defined on $a \le x \le b$. Then if $q(x) < 0$ and $[p(x)\varphi_n\varphi_n']_a^b \le 0$, all the eigenvalues are nonnegative.
7. The nth eigenfunction of a regular Sturm–Liouville problem defined on the interval $a \le x \le b$ has exactly $n - 1$ zeros lying strictly inside the interval.
8. Let two regular Sturm–Liouville problems defined on an interval $a \le x \le b$ be such that $[p(x)\varphi_n\varphi_n']_a^b = 0$ and differ only in their coefficients $p(x)$. Furthermore, let the problem with the coefficient $p_1(x)$ have the eigenvalues $\lambda_1^{(1)}, \lambda_2^{(1)}, \ldots$, and the problem with the coefficient $p_2(x)$ have the eigenvalues $\lambda_1^{(2)}, \lambda_2^{(2)}, \ldots$. Then, if $p_1(x) > p_2(x)$,

$$\lambda_n^{(1)} > \lambda_n^{(2)} \quad \text{for } n = 1, 2, \ldots.$$

9. Let a regular Sturm–Liouville equation with $q(x) < 0$ be defined on an interval $a \le x \le b$ and have boundary conditions such that the first term in the numerator of the Rayleigh quotient in Property 5 is zero. Then reducing the length of the interval $a \le x \le b$ will not reduce the value of any eigenvalue. ■

Remarks about Theorem 8.3

Property 1 ensures that the eigenvalues are distinct ($\lambda_m \ne \lambda_n$ if $m \ne n$), that they are infinite in number, and, because $\lim_{n\to\infty} \lambda_n = \infty$, that there can be no clustering of eigenvalues about a finite limit point. If, for example, the eigenvalues represent the frequencies of vibration of a stretched string of finite length L, this means there is a lowest frequency of vibration, but no upper limit to the frequency of vibration of the string.

Property 2 says that to each distinct eigenvalue of a regular Sturm–Liouville problem there corresponds only one eigenfunction, and it is unique apart from a constant multiplicative factor. Notice that this only applies to *regular* Sturm–Liouville problems, because in periodic Sturm–Liouville problems an eigenvalue has associated with it two linearly independent eigenfunctions. This latter situation occurred in Example 8.19, where the *two* eigenfunctions

$$y_n^{(1)}(x) = \cos(2n\pi x/L) \quad \text{and} \quad y_n^{(2)}(x) = \sin(2n\pi x/L)$$

were seen to correspond to the *single* eigenvalue $\lambda_n = 4n^2\pi^2/L^2$. In such cases there can only be two eigenfunctions to each eigenvalue, because the equation is second order. The scaling of eigenfunctions by a constant is used repeatedly when representing arbitrary functions in terms of series of eigenfunctions.

Property 3 is of fundamental importance because of the part played by orthogonality when developing arbitrary functions in terms of series of eigenfunctions defined over some interval. It is the orthogonality of sine and cosine functions illustrated in Example 8.22 that is used when working with Fourier series.

It will be seen later that the representation (*expansion*) of arbitrary functions in terms of series of eigenfunctions is more general than in terms of power series. This is because, unlike Taylor series whose coefficients are determined by repeated differentiation of the function being expanded, the coefficients in series of eigenfunctions are determined in terms of integrals involving the function. This means that the function can have finite discontinuities at points within its interval of representation and still have an eigenfunction expansion.

Property 4 removes the necessity to check Sturm–Liouville problems for the possibility that negative eigenvalues occur. Had this property been known in advance of Examples 8.18 to 8.21, it would have been unnecessary to have examined the forms of solution corresponding to $\lambda < 0$.

Property 5 is useful when seeking qualitative properties of eigenvalues. The result is not directly useful when trying to determine an eigenvalue because the associated eigenfunction needs to be known. The main use of the Rayleigh quotient arises when it is used in the following rather different form.

Let a function $\Phi(x)$ containing some arbitrary constants $\alpha, \beta, \ldots$ satisfy the *boundary conditions* of a Sturm–Liouville problem. Then with any choice of the arbitrary constants, the Rayleigh quotient

$$\frac{-[p\Phi_n\Phi_n']_a^b + \int_a^b p(\Phi_n')^2\,dx - \int_a^b q\Phi_n^2\,dx}{\int_a^b r\Phi_n^2\,dx} \tag{128}$$

provides an *upper bound* for the value of the smallest eigenvalue of the associated Sturm–Liouville problem. If the arbitrary constants $\alpha, \beta, \ldots$ are chosen to *minimize* this expression, its value becomes the best estimate of the smallest eigenvalue that can be obtained using that approximation. Furthermore, substituting the values of the constants that minimize the Rayleigh quotient into the function $\Phi(x)$ provides a corresponding approximation to the first eigenfunction. The actual value λ_1 is only attained when $\Phi(x) = \varphi_1(x)$.

Property 6, together with Property 4, ensures that under the given conditions the eigenvalues are both real and positive. In corresponding physical problems this result is usually to be expected on an intuitive basis, so the result provides the mathematical justification for making such an assumption on purely physical grounds.

Property 7 provides precise information about the number of zeros of a given eigenfunction within the interval over which it is defined. It is well illustrated by considering Figs. 8.1 showing graphs of Legendre polynomials. These show, for example, that $P_3(x)$ has precisely three zeros in the interval $-1 \leq x \leq 1$, whereas $P_4(x)$ has four zeros. It is important to recognize that these zeros lie strictly *inside* the interval, so that zeros that occur at either end of an interval are *not* counted.

Property 8 means that if in a Sturm–Liouville problem $p(x)$ is associated with a characteristic feature of a physical system, then increasing $p(x)$ increases each eigenvalue of the system. For example, if $p(x)$ is related to the density of a vibrating string, then *increasing* the density while keeping all other parameters constant will *decrease* the frequency of vibration, and increasing the tension will increase the frequency.

Property 9 means that reducing the length of the interval $a \leq x \leq b$ on which a Sturm–Liouville problem is set cannot reduce the values of the eigenvalues. In fact, it usually increases them. This is most easily understood in terms of a vibrating string for which the eigenvalues of the associated Sturm–Liouville problem represent its possible frequencies of vibration (see Chapter 18). In such a case *shortening* the string, while leaving other parameters unchanged, will *increase* the frequency, as any guitarist or violinist knows from experience.

EXAMPLE 8.23

orthogonality and weight functions

An orthogonal system of sine functions The Sturm–Liouville problem considered in Example 8.18, namely

$$y'' + \lambda y = 0 \quad \text{with } y(0) = 0 \quad \text{and} \quad y'(\pi) = 0,$$

is such that $p(x) = 1, q(x) = 0$, and $r(x) = 1$. Its eigenvalues were shown to be $\lambda_n = (2n+1)^2/4$, and its corresponding eigenfunctions were

$$\varphi_n(x) = \sin\frac{(2n+1)x}{2}, \quad n = 0, 1, \ldots.$$

Thus, from Theorem 8.3 (3), the functions $\varphi_n(x)$ are orthogonal over the interval $0 \leq x \leq \pi$ with weight function $r(x) = 1$, and so

$$\int_0^\pi \varphi_m(x)\varphi_n(x)dx = 0 \quad \text{for } m \neq n.$$

The square of the norm is given by

$$\|\varphi_n(x)\|^2 = \left\|\sin\frac{(2n+1)x}{2}\right\|^2 = \int_0^\pi \left(\sin\frac{(2n+1)x}{2}\right)^2 dx = \frac{\pi}{2},$$

so $\|\varphi_n(x)\| = \sqrt{\pi/2}$. ■

EXAMPLE 8.24

Orthogonality of Legendre polynomials When written in Sturm–Liouville form, Legendre's equation becomes

$$[(1-x^2)y']' + \lambda y = 0,$$

and it is defined over the interval $-1 \leq x \leq 1$, with $p(x) = 1 - x^2$, $q(x) = 0$, and $r(x) = 1$. The Legendre polynomial $P_n(x)$ corresponds to $\lambda = n(n+1)$, so from

Theorem 8.3 (3) we see that the Legendre polynomials are orthogonal with respect to the weight function $r(x) = 1$, so that

$$\int_{-1}^{1} P_m(x)P_n(x)dx = 0 \quad \text{for } m \neq n.$$

To determine the norm $\|P_n(x)\|$ we make use of recurrence relation (16),

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0.$$

Replacing n by $n-1$ and substituting for one of the factors $P_n(x)$ in the integral gives

$$\begin{aligned}\|P_n(x)\|^2 &= \int_{-1}^{1} P_n(x)\left\{\left(\frac{2n-1}{n}\right)xP_{n-1}(x) - \frac{n-1}{n}P_{n-2}(x)\right\}dx \\ &= \left(\frac{2n-1}{n}\right)\int_{-1}^{1} xP_{n-1}(x)P_n(x)dx - \left(\frac{n-1}{n}\right)\int_{-1}^{1} P_n(x)P_{n-2}(x)dx \\ &= \left(\frac{2n-1}{n}\right)\int_{-1}^{1} xP_{n-1}(x)P_n(x)dx,\end{aligned}$$

where the second integral has been set equal to zero because of the orthogonality of $P_n(x)$ and $P_{n-2}(x)$. Using the recurrence relation to remove the term $xP_n(x)$ gives

$$\begin{aligned}\|P_n(x)\|^2 &= \left(\frac{2n-1}{n}\right)\int_{-1}^{1} P_{n-1}(x)\left\{\left(\frac{n+1}{2n+1}\right)P_{n+1}(x) + \left(\frac{n}{2n+1}\right)P_{n-1}(x)\right\}dx \\ &= \left(\frac{2n-1}{2n+1}\right)\int_{-1}^{1} [P_{n-1}(x)]^2\,dx,\end{aligned}$$

where the first integral vanishes because of the orthogonality of $P_n(x)$ and $P_{n-1}(x)$.

This has established the recurrence relation for norms

$$\|P_n(x)\|^2 = \left(\frac{2n-1}{2n+1}\right)\|P_{n-1}(x)\|^2.$$

Using this result to relate $\|P_n(x)\|^2$ to $\|P_0(x)\|^2$ and cancelling factors shows that

$$\begin{aligned}\|P_n(x)\|^2 &= \left(\frac{2n-1}{2n+1}\right)\left(\frac{2n-3}{2n-1}\right)\left(\frac{2n-5}{2n-3}\right)\cdots\left(\frac{3}{5}\right)\left(\frac{1}{3}\right)\|P_0(x)\|^2 \\ &= \left(\frac{1}{2n+1}\right)\|P_0(x)\|^2,\end{aligned}$$

but $\|P_0(x)\|^2 = \int_{-1}^{1} 1dx = 2$, so that

$$\|P_n(x)\|^2 = \frac{2}{2n+1}, \quad \text{and} \quad \|P_n(x)\| = \sqrt{\frac{2}{2n+1}} \quad \text{for } n = 0, 1, \ldots.$$ ■

EXAMPLE 8.25

Orthogonality of Bessel functions $J_n(x)$ When written in Sturm–Liouville form, Bessel's equation of order n becomes

$$[xJ_n'(kx)]' + \left(k^2x - \frac{n^2}{x}\right)J_n(kx) = 0,$$

where $p(x) = x$, $q(x) = -n^2/x$, $r(x) = x$, and $\lambda = k^2$.

The orthogonality of Bessel functions over an interval $0 \leq x \leq a$ takes a somewhat different form from that in the previous examples, because the orthogonality is between Bessel functions of the *same* order, but with *different* arguments, rather than between Bessel functions of different orders. If for fixed n the solution $J_n(kx)$ is required to satisfy the boundary condition

$$J_n(ka) = 0,$$

it follows, as in Example 8.21, that the permissible values of k are

$$k_r = j_{n,r}/a, \quad \text{with } r = 1, 2, \ldots,$$

where $j_{n,r}$ is the rth zero of $J_n(x)$, the first few of which are listed in Table 8.1.

Theorem 8.3 (3) then asserts that as the weight function $r(x) = x$, the orthogonality condition is

$$\int_0^a xJ_n\left(\frac{j_{n,r}x}{a}\right)J_n\left(\frac{j_{n,s}x}{a}\right)dx = 0 \quad \text{for } r \neq s.$$

The square of the norm of $J_n(\frac{j_{n,r}x}{a})$ is

$$\left\| J_n\left(\frac{j_{n,r}x}{a}\right)\right\|^2 = \int_0^a x\left[J_n\left(\frac{j_{n,r}x}{a}\right)\right]^2 dx = \frac{a^2}{2}[J_{n+1}(j_{n,r})]^2.$$

A proof of this last result is given in Appendix 2 at the end of the chapter. ■

EXAMPLE 8.26

Orthogonality of Chebyshev polynomials When written in Sturm–Liouville form, the Chebyshev equation for the polynomial $T_n(x)$ of degree n becomes

$$[(1-x^2)^{1/2}y']' + n^2(1-x^2)^{-1/2}y = 0.$$

As the weight function is $(1-x^2)^{-1/2}$, the orthogonality relation becomes

$$\int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}dx = 0 \quad \text{for } m \neq n.$$

The square of the norm of $T_n(x)$ is given by

$$\|T_n(x)\|^2 = \int_{-1}^{1} \frac{[T_n(x)]^2}{\sqrt{1-x^2}}dx$$

where $\|T_0(x)\|^2 = \pi$ and $\|T_n(x)\|^2 = \pi/2$ for $n = 1, 2, \ldots$. As it is inappropriate to include the proof of this result here, an outline proof is given in Exercise 31 at the end of the section.

Accounts of Sturm–Liouville systems are to be found in references [3.3] and [3.4] and in Chapter 5 of reference [3.7]. ■

Summary

The important idea of Sturm–Liouville systems was introduced, their relationship to eigenvalues and eigenfunctions was explained, and it was shown that the solutions of such systems comprise a system of functions that are orthogonal with respect to a suitable weight function. The examples of Sturm–Liouville systems that were given included trigonometric, Legendre, Chebyshev, and Bessel functions. Infinite sets of functions like these represent generalizations to an infinite dimensional space of the elementary notion of the orthogonality of vectors in the three-dimensional Euclidean space. The significance of the orthogonality of eigenfunctions will become clear later when arbitrary functions are expanded in terms of eigenfunctions.

EXERCISES 8.10

In Exercises 1 through 4, reduce the differential equation to Sturm–Liouville form by the method used when reducing equation (121) to the form in (122).

1. $xy'' + (1 - x)y' + \lambda y = 0.$

2. $y'' - 2xy' + \lambda y = 0.$

3. $(1 - x^2)y'' - xy' + \lambda y = 0.$

4. $(1 - x^2)^2 y'' - 2x(1 - x^2)y' + [\lambda(1 - x^2) - m^2]y = 0.$

In Equations 5 through 14 find the eigenvalues and eigenfunctions of the differential equation.

5. $y'' + \lambda y = 0, \quad y(0) = 0, \quad y(L) = 0.$

6. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y'(L) = 0.$

7. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y(1) = 0.$

8. $y'' + \lambda y = 0, \quad y(0) = 0, \quad y'(2\pi) = 0.$

9. $y'' + \lambda y = 0$, $y(0) = 0$, $y'(1) - 2y(1) = 0$. Find numerical estimates for the first two eigenvalues.

10. $y'' + \lambda y = 0$, $y(0) = 0$, $y'(1) + y(1) = 0$. Find numerical estimates for the first two eigenvalues.

11. $y'' + \lambda y = 0, \quad y(-1) = y(1), \quad y'(-1) = y'(1).$

12. $y'' + \lambda y = 0, \quad y(0) = y(1), \quad y'(0) = y'(1).$

13. $x^2y'' + xy' + k^2y = 0, \quad y(1) = 0, \quad y(4) = 0.$ (Hint: This is a Cauchy–Euler equation)

14. $x^2y'' + xy' + 9k^2y = 0, \quad y(1) = 0, \quad y'(2) = 0.$ (Hint: This is a Cauchy–Euler equation)

In Exercises 15 through 18, verify that the sets of functions are orthogonal over their stated intervals with the weight function $r(x) = 1$, and find their norms.

15. $\varphi_n(x) = \sin\left(\dfrac{n\pi x}{L}\right), \quad n = 1, 2, \ldots \ (0 \le x \le L).$

16. $\varphi_n(x) = \cos\left(\dfrac{(2n-1)\pi x}{2}\right), \quad n = 1, 2, \ldots \ (0 \le x \le 1).$

17. $\varphi_n(x) = \cos\left(\dfrac{n\pi x}{L}\right), \quad n = 1, 2, \ldots \ (0 \le x \le L).$

18. $\varphi_n(x) = \sin\left(\dfrac{(2n-1)\pi x}{4}\right), \ n = 1, 2, \ldots \ (0 \le x \le 2\pi).$

19.* It is known from Example 8.18 that the Sturm–Liouville problem

$$y'' + \lambda y = 0 \quad \text{with } y(0) = 0, \ y'(\pi) = 0$$

has for its first eigenvalue $\lambda_1 = 1/4$, and that the corresponding eigenfunction is $\varphi_1(x) = \sin x/2$. Verify that the function $\Phi(x) = x(2\pi - x)$ satisfies the boundary conditions for y. By using this expression in the form of the Rayleigh quotient given in (128), find the corresponding upper bound for λ_1 and compare it with the exact value. Why is it that replacing $\Phi(x)$ by $\Phi(x) = Cx(2\pi - x)$, where C is any nonzero constant, leaves the estimate of the upper bound unchanged?

20.* Perform the calculation required in Exercise 19 using the function $\Phi(x) = x^2(1 - \frac{2x}{3\pi})$, after first showing that $\Phi(x)$ satisfies the boundary conditions. Compare the value of the upper bound so obtained with the exact value $\lambda_1 = 1/4$. Suggest a reason why this approximation is not likely to yield a better lower bound than the one obtained using the function $\Phi(x)$ in Exercise 19.

21.* The Sturm–Liouville form of Bessel's equation of order 1 is

$$[xy']' + \left(k^2x - \frac{1}{x}\right)y = 0,$$

where $p(x) = x$, $q(x) = -1/x$, $r(x) = x$, and $\lambda = k^2$. The bounded solution of this equation on the interval $0 \le x \le 1$ subject to the condition $y(1) = 0$ is $y(x) = J_1(j_{1,1}x)$, where from Table 8.1 $j_{1,1} = 3.8317$ is the first zero of $J_1(x)$. The inverted parabola $\Phi(x) = x(1 - x)$ provides a reasonable approximation to the shape of the required Bessel function for $0 \le x \le 1$. Use this expression in (128) to obtain an upper bound for the first eigenvalue λ_1 of the equation, and using the fact that $\lambda_1 = j_{1,1}^2$ find an upper bound for $j_{1,1}$. Compare this estimate with the correct result.

22.* The Sturm–Liouville form of Bessel's equation of order 2 is

$$[xy']' + \left(k^2x - \frac{4}{x}\right)y = 0,$$

where $p(x) = x$, $q(x) = -4/x$, $r(x) = x$, and $\lambda = k^2$. The solution of this equation that is bounded on the interval $0 \le x \le 1$ and subject to the condition $y(1) = 0$ is $y(x) = J_2(j_{2,1}x)$, where from Table 8.1 $j_{2,1} = 5.1316$ is the first zero of $J_2(x)$. Use the approximation $\Phi(x) = x(1 - x)$ to obtain an upper bound for the first eigenvalue of the equation, and using the fact that $\lambda_1 = j_{2,1}^2$, find an upper bound for $j_{2,1}$. Compare this estimate with the correct value.

23. The differential equation

$$L[y] = P(x)y'' + Q(x)y' + R(x)y = 0$$

has associated with it the **adjoint differential equation** defined by

$$M[w] = [P(x)w]'' - [Q(x)w]' + R(x)w = 0.$$

A differential equation is said to be **self-adjoint** if the differential equation and its adjoint are of the same form. When this occurs, the differential operator common to both equations is also said to be self-adjoint.

(a) Show that Bessel's equation of order ν

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0$$

is not self-adjoint.

(b) Find the value of α that makes the following equation self-adjoint

$$(\alpha \sin x)y'' + (\cos x)y' + 2y = 0.$$

24. Show that Legendre's equation

$$(1 - x^2)y'' - 2xy' - \lambda y = 0$$

is self-adjoint.

25. Show that Bessel's equation of order n in the form

$$x^2y'' + xy' - (x^2 - n^2)y = 0$$

is not self-adjoint, but that it becomes so when multiplied by $1/x$.

26. Show that the Hermite equation in the form

$$y'' - 2xy' + \lambda y = 0$$

is not self-adjoint, but that it becomes so when multiplied by $\exp[-x^2]$.

27. Show that the Chebyshev equation in the form

$$(1 - x^2)y'' - xy' + \lambda y = 0$$

is not self-adjoint, but that it becomes self-adjoint when multiplied by $(1 - x^2)^{-1/2}$.

28.* Let $u(x)$ and $v(x)$ be any two solutions of

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y = 0$$

defined over the interval $a \le x \le b$. Prove **Abel's identity**

$$p(x)[u(x)v'(x) - u'(x)v(x)] = \text{constant}$$

for all x in the interval. As $p(x) \neq 0$ in regular Sturm–Liouville problems, what conclusion can be drawn from Abel's identity if (a) the constant is not zero and (b) the constant is zero?

(Hint: Multiply the equations for u and v by suitable factors, subtract them, and integrate the resulting equation over the interval $a \le t \le x$.)

29.* The Chebyshev polynomial $T_n(x)$ can be defined as

$$T_n(x) = \cos(n \arccos x), \quad n = 0, 1, \ldots.$$

Verify this by showing that this definition of $T_n(x)$ satisfies the Chebyshev differential equation

$$(1 - x^2)y'' - xy' + n^2y = 0.$$

30.* Let $y = T_n(x) = \cos(n \arccos x)$ and set $x = \cos\theta$. Use the fact that $y(\theta)$ satisfies the differential equation

$$\frac{d^2y}{d\theta^2} + n^2y = 0$$

together with a change of variable back from θ to x to show that this definition of $T_n(x)$ satisfies the Chebyshev equation

$$(1 - x^2)y'' - xy' + n^2y = 0.$$

31.* Show that if $y_n(\theta) = \cos n\theta$ then

$$\int_0^\pi [y_n(\theta)]^2 d\theta = \begin{cases} \pi, & n = 0 \\ \frac{1}{2}\pi, & n \ge 1 \end{cases}.$$

By changing back from the variable θ to x, where $x = \cos\theta$ and using the definition of $T_n(x)$ in Problem 30, show that the square of the norm of $T_n(x)$ is given by

$$\|T_n(x)\|^2 = \int_{-1}^1 \frac{[T_n(x)]^2}{\sqrt{1 - x^2}}dx = \begin{cases} \pi, & n = 0 \\ \frac{1}{2}\pi, & n \ge 1 \end{cases}.$$

8.11 Eigenfunction Expansions and Completeness

The orthogonality of a set of functions $\varphi_0(x), \varphi_1(x), \ldots$ over the interval $a \le x \le b$ with respect to a weight function $r(x)$ allows them to be used to expand (represent) a function $f(x)$ over that same interval in terms of the functions $\varphi_i(x)$ by expressing it as the series

$$f(x) = \sum_{m=0}^{\infty} a_m\varphi_m(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots, \tag{129}$$

where $a_0, a_1, \ldots$ are constants called the **coefficients** of the expansion.

The representation of functions in this manner is used in approximation theory, in numerical analysis, and in the solution of partial differential equations by the method of *separation of variables* to be described later (see Chapter 18). A series eigenfunction expansions

such as (129) is called a **generalized Fourier series** representation of $f(x)$ or, when the functions $\varphi_n(x)$ are eigenfunctions, an **eigenfunction expansion** of $f(x)$.

To see how the coefficients a_m in (129) are derived for a specific function $f(x)$, it is necessary to recall that

$$\int_a^b r(x)\varphi_m(x)\varphi_n(x)dx = 0, \quad m \neq n, \tag{130}$$

and

$$\|\varphi_n(x)\|^2 = \int_a^b r(x)[\varphi_n(x)]^2\,dx. \tag{131}$$

If the expansion (129) is multiplied by $r(x)\varphi_n(x)$ and the result is integrated over the interval $a \le x \le b$, the orthogonality condition (130) causes every term on the right for which $m \neq n$ to vanish, leaving only the term involving a_n, so using (131) enables the result to be written

$$\int_a^b r(x)\varphi_n(x)f(x)dx = a_n\int_a^b r(x)[\varphi_n(x)]^2dx = a_n\|\varphi_n(x)\|^2.$$

This has established that the coefficients a_n are given by the formula

$$a_n = \frac{\int_a^b r(x)\varphi_n(x)f(x)dx}{\|\varphi_n(x)\|^2}, \quad n = 0, 1, \ldots. \tag{132}$$

The term-by-term integration of series (129) leading to (132) requires justification, and this follows when the series is uniformly convergent.

Summary of Main Sets of Orthogonal Functions

1. Fourier series (see Chapter 9)

Interval of definition	$-\pi \le x \le \pi$
Set of functions	$\{1, \cos nx, \sin nx\}$, $n = 1, 2, \ldots$
Weight	$r(x) = 1$
Orthogonality	$\int_{-\pi}^{\pi} \sin mx \sin nx\,dx = 0,\ m \neq n$
	$\int_{-\pi}^{\pi} \sin mx \cos nx\,dx = 0,$
	$\int_{-\pi}^{\pi} \cos mx \cos nx\,dx = 0,\ m \neq n$
	$\int_{-\pi}^{\pi} 1 \cdot \sin mx\,dx = 0$
	$\int_{-\pi}^{\pi} 1 \cdot \cos mx\,dx = 0$
Norms	$\|1\|^2 = 2\pi,\ \|\cos nx\|^2 = \pi,\ \|\sin nx\|^2 = \pi$

2. Legendre polynomials

Interval of definition	$-1 \le x \le 1$
Set of functions	$P_0(x) = 1,\ P_1(x) = x,\ P_2(x) = \frac{1}{2}(3x^2 - 1), \ldots$
Recurrence relation	$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0$
Weight	$r(x) = 1$
Orthogonality	$\int_{-1}^{1} P_m(x)P_n(x)dx = 0,\ m \ne n$
Norm	$\lVert P_n(x)\rVert^2 = \dfrac{2}{2n+1},\ n = 0, 1, \ldots.$

3. Bessel functions

Interval of definition	$0 \le x \le a$
Set of functions	There is a set of orthogonal functions for each fixed n: $J_n(j_{n,r}x/a),\ r = 1, 2, \ldots,$ with $j_{n,r}$ the nth zero of $J_n(x)$ (see Table 8.1)
Weight	$r(x) = x$
Orthogonality	$\int_0^a xJ_n(j_{n,r}x/a)J_n(j_{n,s}x/a)dx = 0,\ r \ne s$
Norm	$\lVert J_n(j_{n,r}x/a)\rVert^2 = \frac{1}{2}a^2[J_{n+1}(j_{n,r})]^2,\ r = 1, 2, \ldots$

4. Chebyshev polynomials

Interval of definition	$-1 \le x \le 1$
Set of functions	$T_0(x) = 1,\ T_1(x) = x,\ T_2(x) = 2x^2 - 1, \ldots$
Recurrence relation	$T_{n+1}(x) - 2xT_n(x) + T_{n-1}(x) = 0$
Weight	$(1 - x^2)^{-1/2}$
Orthogonality	$\int_{-1}^{1} \dfrac{T_m(x)T_n(x)}{\sqrt{1-x^2}}dx = 0,\ m \ne n$
Norms	$\lVert T_0(x)\rVert^2 = \pi,\ \lVert T_n(x)\rVert^2 = \frac{1}{2}\pi,\quad n = 1, 2, \ldots.$

(See Exercises 30 and 31 in Exercise Set 18.10 for the derivation of the norms.)

EXAMPLE 8.27

a first example of a Fourier series

A Fourier series Example 8.22 established the orthogonality of the set of functions

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$$

over the interval $-\pi \le x \le \pi$ with weight $r(x) = 1$. It is left as a simple exercise to verify that these functions are the eigenfunctions of the Sturm–Liouville problem

$$y'' + \lambda y = 0, \quad y(-\pi) = y(\pi) = 0.$$

The **Fourier series** for a function $f(x)$ is

$$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

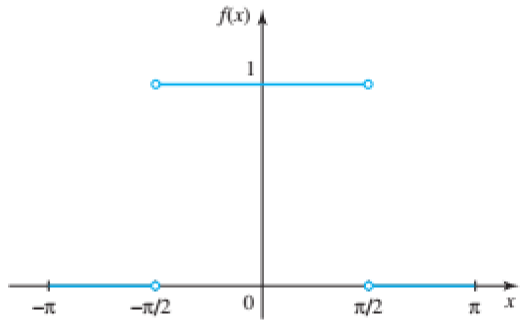


FIGURE 8.13 The rectangular pulse.

where from (132), the **Fourier coefficients** are

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx, \quad a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx, \quad n = 1, 2, \ldots.$$

The formulas for the a_n and b_n are called the **Euler formulas** for the Fourier coefficients.

In anticipation of Chapter 9, let us use these results to find the Fourier series of the (discontinuous) rectangular pulse function

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

shown in Fig. 8.13.

The discontinuities in $f(x)$ cause no problem when deriving the coefficients a_n and b_n because integrals of finite discontinuous functions are well defined:

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} 1\,dx = \frac{1}{2}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\cos nx\,dx = \frac{2}{n\pi}\sin\left(\frac{1}{2}n\pi\right) = \begin{cases} 0 & \text{if } n \text{ is even} \\ \pm\frac{2}{n\pi} & \text{if } n \text{ is odd.} \end{cases}$$

A similar calculation shows that

$$b_n = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\sin nx\,dx = \left[\frac{-1}{n\pi}\cos nx\right]_{-\pi/2}^{\pi/2} = 0, \quad n = 1, 2, \ldots.$$

Substituting for the coefficients in the Fourier series gives

$$f(x) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\cos(2n-1)x}{2n-1},$$

and so

$$f(x) = \frac{1}{2} + \frac{2}{\pi}\left(\cos x - \frac{\cos 3x}{3} + \frac{\cos 5x}{5} - \cdots\right), \quad -\pi \le x \le \pi.$$

Notice that although $f(x)$ is discontinuous at $x = \pm\pi/2$, the Fourier series is defined at these points and has the value 1/2. ■

This example illustrates the fact that a Fourier series expansion (and indeed any eigenfunction expansion) of $f(x)$ is defined for all x in its interval of definition, including points where $f(x)$ is discontinuous, or not even defined. Because of this it is necessary to question the use of the equality sign in (129) and to reinterpret its meaning at points of discontinuity of $f(x)$. More will be said about this in Chapter 9 in connection with Fourier series.

Some comments will be offered later about the convergence of eigenfunction expansions in general, and their behavior at points of discontinuity of $f(x)$ when the completeness of sets of orthogonal functions is discussed.

EXAMPLE 8.28

a Fourier–Legendre expansion

A Fourier–Legendre expansion The expansion of a function $f(x)$ in terms of Legendre polynomials $P_n(x)$ over the interval $-1 \le x \le 1$ is called a **Fourier–Legendre expansion**, and it takes the form

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x) = a_0 + a_1 P_1(x) + \cdots. \tag{133}$$

From (135) the coefficients a_n are determined by

$$a_n = \frac{\int_a^b r(x)\varphi_n(x) f(x)dx}{\|\varphi_n(x)\|^2} = \left[\frac{2n+1}{2}\right] \int_{-1}^{1} f(x)P_n(x)dx, \quad n = 0, 1, \ldots.$$

As any polynomial of degree m can be expressed as a linear combination of $P_0(x), P_1(x), \ldots, P_m(x)$, it follows from the orthogonality condition that

$$\int_{-1}^{1} x^m P_n(x)dx = 0 \text{ for } n > m.$$

The Fourier–Legendre expansion of the discontinuous function

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

is determined as follows. From (133),

$$a_n = \left(\frac{2n+1}{2}\right) \int_{-1}^{1} f(x)P_n(x)dx = \left(\frac{2n+1}{2}\right) \int_{0}^{1} P_n(x)dx. \tag{134}$$

If we substitute for $P_n(x)$, it then follows that the first few coefficients in the expansion are

$$a_0 = \frac{1}{2}, a_1 = \frac{3}{4}, a_2 = 0, a_3 = -\frac{7}{16}, \ldots,$$

so the required expansion is

$$f(x) = \frac{1}{2}P_0(x) + \frac{3}{4}P_1(x) - \frac{7}{16}P_3(x) + \cdots.$$

Here also this Fourier-Legendre expansion attributes a value to $f(x)$ at its point of discontinuity at $x = 0$, and a closer examination shows that the value determined by the expansion is 1/2. ■

EXAMPLE 8.29

a Fourier–Bessel expansion

Fourier–Bessel expansions A function $f(x)$ can be expanded over the interval $0 \le x \le a$ in terms of the Bessel function J_n, with n fixed, to obtain a **Fourier–Bessel** expansion of the form

$$f(x) = \sum_{r=1}^{\infty} a_r J_n(j_{n,r}x/a) = a_1 J_n(j_{n,1}x/a) + a_2 J_n(j_{n,2}x/a) + \cdots, \tag{135}$$

where

$$a_r = \left(\frac{2}{a^2}\right) \frac{\int_0^a J_n(j_{n,r}x/a) f(x)dx}{[J_{n+1}(j_{n,r})]^2} \tag{136}$$

An expansion of this type will be used in Chapter 18 when solving the oscillations of a circular membrane, such as the membrane covering a circular drum head. ■

EXAMPLE 8.30

a Fourier–Chebyshev expansion

Fourier–Chebyshev expansions The **Fourier–Chebyshev expansion** of a function $f(x)$ over the interval $-1 \le x \le 1$ takes the form

$$f(x) = \sum_{n=0}^{\infty} a_n T_n(x) = a_0 T_0(x) + a_1 T_1(x) + \cdots, \tag{137}$$

where

$$a_n = \frac{\int_{-1}^{1} \dfrac{f(x)T_n(x)}{\sqrt{1-x^2}} dx}{\| T_n(x) \|^2}, \tag{138}$$

with

$$\| T_0(x) \|^2 = \pi \quad \text{and} \quad \| T_n(x) \|^2 = \frac{1}{2}\pi, \quad n = 1, 2, \ldots .$$

Any polynomial of degree m can be expressed as a linear combination of $T_0(x)$, $T_1(x), \ldots, T_m(x)$, so it follows from the orthogonality conditions that

$$\int_{-1}^{1} \frac{x^m T_n(x)}{\sqrt{1-x^2}} dx = 0 \quad \text{for } n > m.$$ ■

It is now necessary to comment on the interpretation of the equality sign in (129) at points where $f(x)$ is discontinuous. For expansions in terms of orthogonal functions to be useful, they must be able to represent the class of functions that occur in practical applications. This means that an orthogonal set of functions defined over an interval $a \le x \le b$ must always be able to be used to expand functions that are piecewise continuous and differentiable at all but a finite number of points in the interval. For conciseness we will denote this set of functions by PC. In addition, the set of orthogonal functions must be sufficiently rich in functions that there is no function of practical importance that cannot be expanded in this manner.

completeness and convergence

Orthogonal (and orthonormal) sets of functions that have this property are said to be **complete**, and the ones introduced so far can all be shown to have this property of completeness. As sets of orthogonal functions are required to expand both continuous and piecewise continuous functions that belong to class PC, the convergence of these expansions must of necessity be more general in nature than ordinary convergence. It is this more general form of convergence, which will be introduced shortly, that will permit the equality sign in (129) to be interpreted in a special sense at points where $f(x)$ is discontinuous.

The special type of convergence we now introduce is called **convergence in the norm**, **mean-square convergence**, or **$\mathbf{L}^2$ convergence**. This form of convergence is defined by requiring that if a sequence of functions $f_1(x)$, $f_2(x)$, ... *converges in the mean* to a function $f(x)$, then

$$\lim_{n\to\infty} \| f_n(x) - f(x)\| = 0, \tag{139}$$

or, more explicitly,

$$\lim_{n\to\infty} \int_a^b r(x)[f_n(x) - f(x)]^2\, dx = 0. \tag{140}$$

When interpreting (139) as (140) it is convenient to omit the square root in the definition of the norm, as this simplifies analysis and does not influence the limit.

The sequence of functions $f_n(x)$ in this definition can be taken to be the nth partial sum of the eigenfunction expansion (129),

$$f_n(x) = \sum_{m=0}^{n} a_m\varphi_m(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots, \tag{141}$$

where from now on we will assume $\varphi_0(x), \varphi_1(x), \ldots$ to be an orthonormal set of functions so that $\|\varphi_n(x)\|^2 = 1, n = 0, 1, \ldots$. Such an orthonormal set of functions will be complete with respect to the functions $f(x)$ in C if every function in PC can be approximated by (141). Convergence in the norm and ordinary convergence are the same everywhere a function is continuous and differentiable.

We now state without proof the fundamental eigenfunction expansion theorem.

THEOREM 8.4

a fundamental eigenfunction expansion theorem

Eigenfunction expansion theorem Let $f(x)$ and $f'(x)$ have at most a finite number of jump discontinuities in the interval $a \le x \le b$. Then the eigenfunction expansion (129) converges in the mean to $f(x)$ at every point of continuity of $f(x)$ inside this interval, and to the value $\frac{1}{2}[f(c-) + f(c+)]$ at any point c where $f(x)$ is discontinuous. ■

This convergence property has already been demonstrated in Example 8.27, where the Fourier series converged to the value $1/2$ at the points where the function was discontinuous. Figure 8.14 shows the result in the general case.

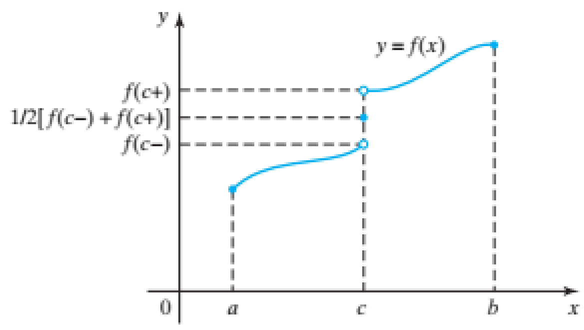


FIGURE 8.14 Convergence of an eigenfunction expansion at a point of discontinuity.

To develop the concept of completeness a little further, we substitute (129) into (140) to obtain

$$\int_a^b r(x)[f_n(x) - f(x)]^2 dx = \int_a^b r(x)[f_n(x)]^2 dx - 2\int_a^b r(x)f(x)f_n(x)dx$$
$$+ \int_a^b r(x)[f(x)]^2 dx = \int_a^b r(x)\left[\sum_{s=0}^{n} a_s\varphi_s(x)\right]^2 dx$$
$$- 2\sum_{s=0}^{n} a_s \int_a^b r(x)f(x)\varphi_s(x)dx + \int_a^b r(x)[f(x)]^2 dx.$$

The orthogonality property of the set of eigenfunctions $\varphi_s(x)$ reduces the first integral on the right to $\sum_{s=0}^{n} a_s^2$, while the definition of a_s shows that the second term on the right can be written $-2\sum_{s=0}^{n} a_s^2$, so the result becomes

$$\int_a^b r(x)[f_n(x) - f(x)]^2 dx = -\sum_{s=0}^{n} a_s^2 + \int_a^b r(x)[f(x)]^2 dx.$$

The integrands of both integrals are nonnegative, and the integral on the right is $\| f(x)\|^2$, so we have established the inequality

$$\sum_{s=0}^{n} a_s^2 \le \int_a^b r(x)[f(x)]^2 dx = \| f(x)\|^2 \quad \text{for all } n \ge 0. \tag{142}$$

Bessel's inequality

This result is called **Bessel's inequality**, and it shows that the sum $\sum_{s=0}^{n} a_s^2$ has the upper bound $\| f(x)\|^2$ as $n \to \infty$. As the terms of the series are nonnegative, the series increases as n increases, so it follows that $\sum_{s=0}^{n} a_s^2$ converges as $n \to \infty$.

If the system of orthonormal functions $\varphi_s(x)$ is complete, result (139) must be true for every function $f(x)$ in the class PC, so that then $\lim_{n\to\infty} \sum_{s=0}^{n} a_s^2 = \| f(x)\|^2$. Consequently, for complete orthonormal systems of functions

$$\sum_{s=0}^{\infty} a_s^2 = \| f(x)\|^2 = \int_a^b r(x)[f(x)]^2 dx. \tag{143}$$

Parseval relation

This result is called the **Parseval relation**.

THEOREM 8.5

Completeness of orthonormal systems Let $\varphi_0(x), \varphi_1(x), \ldots$ be a complete orthonormal set of functions with respect to the set C to which the functions $f(x)$ belong. Then the only continuous function in C that is orthogonal to every function $\varphi_n(x)$ is the zero function $f(x) \equiv 0$. Furthermore, if the restriction of continuity is removed, the only functions that can be orthogonal to every function in the orthonormal set are those with zero norm.

Proof In the first case the vanishing of the norm of $f(x)$ implies that $f(x) \equiv 0$. In the second case, the orthogonality of a function with respect to every eigenfunction implies that the function must be degenerate, and although not identically zero, must have a zero norm. ■

See Chapters 2 and 5 of reference [3.7] for information about eigenfunction expansions and orthonormal sets of functions.

Summary

Eigenfunction expansions have been introduced, and the most important sets of orthogonal functions summarized together with their intervals of definition, weight functions, and orthogonality relationships. Mean-square convergence has been defined and the fundamental eigenfunction theorem stated, and the notion of completeness of systems of orthogonal functions has been explained and related to the Parseval relation.

Appendix 1 (Proofs of Theorem 8.3)

The study of Sturm–Liouville problems is made more concise by the introduction of the notion of a **differential operator *L*** defined as

$$L \equiv \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + q(x), \tag{144}$$

with the understanding that if y is a suitably differentiable function,

$$L[y] \equiv \frac{d}{dx}\left[p(x)\frac{d}{dx}\right]y(x) + q(x)y(x). \tag{145}$$

Differential operators, of which L is a special case, have the property that when they operate on a function y they produce another function $L[y]$. For example, if

$$L \equiv \frac{d}{dx}\left[x\frac{d}{dx}\right] + 2,$$

and $y(x) = e^{-x}$, then

$$L[e^{-x}] = \frac{d}{dx}\left[x\frac{d[e^{-x}]}{dx}\right] + 2e^{-x} = \frac{d}{dx}[-xe^{-x}] + 2e^{-x} = (1+x)e^{-x}.$$

In terms of the differential operator L in (144), the Sturm–Liouville equation (122) with eigenvalue λ and corresponding eigenfunction φ becomes

$$L[\varphi] + \lambda r\varphi = 0, \tag{146}$$

where φ satisfies suitable boundary conditions.

The proof of the results of Theorem 8.3 that can be given here is simplified by appeal to the following theorem, which is important in its own right.

THEOREM 8.6 **One-dimensional form of Green's theorem** Let L be the linear operator

$$L \equiv \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + q(x),$$

and, let u, v be any two twice differentiable functions defined on the interval $a \le x \le b$. Then,

(i)

$$\int_a^b uL[v]dx = [p(x)u(x)v'(x)]_a^b - \int_a^b pu'v'\,dx + \int_a^b quv\,dx$$

and

(ii)

$$\int_a^b \{uL[v] - vL[u]\}dx = [p(x)\{u(x)v'(x) - v(x)u'(x)\}]_a^b,$$

called the **Lagrange identity**. Furthermore, if u and v satisfy the boundary conditions

$$A_1\phi(a) + A_2\phi'(a) = 0 \quad \text{and} \quad B_1\phi(b) + B_2\phi'(b) = 0,$$

where ϕ may be either u or v, then

(iii)

$$\int_a^b \{uL[v] - vL[u]\}dx = 0.$$

Proof Result (i) is the one-dimensional form of **Green's first theorem**, and result (ii) is the one-dimensional form of **Green's second theorem**. The three-dimensional forms of these theorems are derived in Chapter 12, Section 12.2. Result (iii) is the consequence of Green's second theorem when u and v satisfy the stated boundary conditions at the ends of the interval $a \leq x \leq b$.

The proof proceeds as follows. Differentiation of the product $u(pv')$ gives

$$[u(pv')]' = u(pv')' + u'(pv'),$$

so

$$u(pv')' = [puv']' - pu'v'.$$

Recalling the definition of L, we can write

$$uL[v] = [puv']' - pu'v' + quv,$$

so integrating over the interval $a \leq x \leq b$ gives

$$\int_a^b uL[v]dx = [p(x)u(x)v'(x)]_a^b - \int_a^b pu'v'dx + \int_a^b quvdx,$$

which is result (i).

Result (ii) follows if we interchange u and v in (i) and subtract the result from (i) to obtain

$$\int_a^b \{uL[v] - vL[u]\}dx = [p(x)\{u(x)v'(x) - v(x)u'(x)\}]_a^b.$$

Result (iii) follows from (ii) if we notice that, provided $A_2 \neq 0$, it follows from the boundary conditions at $x = a$ that

$$u'(a) = -(A_1/A_2)u(a) \quad \text{and} \quad v'(a) = -(A_1/A_2)v(a),$$

so

$$[p(uv' - vu')]_{x=a} = -(A_1/A_2)p(a)u(a)v(a) + (A_1/A_2)p(a)u(a)v(a) = 0,$$

and a similar argument shows that, provided $B_2 \neq 0$,

$$[p(uv' - vu')]_{x=b} = 0.$$

Thus, $[p(uv' - vu')]_a^b = 0$, reducing result (ii) to

$$\int_a^b \{uL[v] - vL[u]\}dx = 0,$$

which is result (iii).

Result (iii) is obviously true if the boundary conditions simplify to

$$\phi(a) = 0 \text{ and } \phi(b) = 0 \quad \text{or to } \phi'(a) = 0 \quad \text{and} \quad \phi'(b) = 0,$$

and the modification to the proof needed to show that the result remains true if A_2 and/or B_2 is zero is left as an exercise. ■

JOSEPH LOUIS LAGRANGE (1736–1813)
Lagrange was born in Turin of French extraction and after working in Berlin for twenty years moved to Paris. His many fundamental contributions to mathematics have led to his being regarded as one of the most outstanding mathematicians of his time. He made contributions to algebra, calculus, differential equations, the calculus of variations, and also to mechanics.

We now prove the results in Theorem 8.3 that are straightforward, and refer to the references at the end of the chapter for details of the way in which the more complicated results can be established.

Property 1. The proof of this property is difficult and so will be omitted, but Examples 8.18 to 8.21 illustrate the existence of an ordered sequence of eigenvalues in specific cases.

Property 2. In a regular Sturm–Liouville problem suppose, if possible, that φ and ψ are eigenfunctions corresponding to the single eigenvalue λ. Then each of these functions satisfies the Sturm–Liouville equation, while φ and ψ both satisfy the boundary conditions at $x = a$ so that

$$A_1\varphi(a) + A_2\varphi'(a) = 0 \text{ and } A_1\psi(a) + A_2\psi'(a) = 0.$$

This pair of equations can be considered to determine A_1 and A_2 in terms of φ and ψ at $x = a$. The equations are homogeneous, so there can only be a nontrivial solution for A_1 and A_2 if the determinant of coefficients $W = \varphi(a)\psi'(a) - \varphi'(a)\psi(a)$ vanishes, but this determinant is the Wronskian of the solutions and can only vanish if φ is proportional to ψ, so the result is established.

Property 3. Let φ_m and φ_n be eigenfunctions corresponding to the two distinct eigenvalues λ_m and λ_n of the Sturm–Liouville problem

$$L[y] + \lambda r y = 0$$

defined on $a \le x \le b$ and satisfying homogeneous boundary conditions of the type given in (127). Then it follows that

$$L[\varphi_m] + \lambda_m r\varphi_m = 0 \text{ and } L[\varphi_n] + \lambda_n r\varphi_n = 0.$$

Multiplying the first equation by φ_n and the second by φ_m, subtracting the results, and integrating over the interval $a \le x \le b$ gives

$$\int_a^b \{\varphi_m L[\varphi_n] - \varphi_n L[\varphi_m]\}dx + (\lambda_n - \lambda_m)\int_a^b r\varphi_m\varphi_n\, dx = 0.$$

The first integral vanishes because of the result of Theorem 8.4 (iii), so

$$(\lambda_n - \lambda_m)\int_a^b r\varphi_m\varphi_n dx = 0.$$

The result now follows because $\lambda_n \neq \lambda_m$.

Property 4. The proof is by contradiction. Suppose, if possible, that $\lambda = \alpha + i\beta$ is a complex eigenvalue associated with the complex eigenfunction $\Phi = \varphi + i\psi$. Then as Φ and λ satisfy the Sturm–Liouville equation, we have

$$[p(\varphi + i\psi)']' + [q + (\alpha + i\beta)r](\varphi + i\psi) = 0.$$

This can be written

$$[p\varphi']' + q\varphi + \alpha\varphi r - \beta\psi r + i\{[p\psi']' + q\psi + \beta\varphi r + \alpha\psi r\} = 0.$$

For this to be true, both real and imaginary parts of the equation must vanish, so

$$[p\varphi']' + q\varphi + \alpha\varphi r - \beta\psi r = 0 \quad \text{and} \quad [p\psi']' + q\psi + \beta\varphi r + \alpha\psi r = 0.$$

Multiplying the second equation by i, subtracting it from the first equation, and collecting terms gives

$$[p(\varphi - i\psi)']' + [q + (\alpha - i\beta)r](\varphi - i\psi) = 0,$$

showing that $\overline{\Phi} = \varphi - i\psi$ is an eigenfunction and $\overline{\lambda} = \alpha - i\beta$ is an eigenvalue. As Φ and $\overline{\Phi}$ are linearly independent eigenfunctions, it follows from Theorem 8.3 (3) that

$$\int_a^b r\Phi\overline{\Phi}dx = \int_a^b r(\varphi^2 + \psi^2)dx = 0,$$

but this is impossible because by hypothesis $r(x) \geq 0$ and $\varphi^2 + \psi^2 > 0$. Consequently the assumption that an eigenvalue can be complex is false.

Property 5. Let λ_n be an eigenvalue and φ_n be the corresponding eigenfunction of the Sturm–Liouville equation

$$L[\varphi_n] + \lambda_n r\varphi_n = 0.$$

Multiplication of this equation by φ_n, followed by integration over the interval $a \leq x \leq b$, gives

$$\int_a^b \varphi_n L[\varphi_n]dx + \lambda_n\int_a^b r\varphi_n^2 dx = 0.$$

An application of Theorem 8.4 (i) with $u = v = \varphi_n$ then gives the result

$$\lambda_n = \frac{-[p\varphi_n\varphi_n']_a^b + \int_a^b p(\varphi_n')^2dx - \int_a^b r\varphi_n^2dx}{\int_a^b r\varphi_n^2dx}.$$

Property 6. This follows directly from Property 5 when $q(x) < 0$ and the condition $[p\varphi_n\varphi_n']_a^b \leq 0$ is satisfied.

Property 7. We offer no proof of this result, though as already remarked it is well illustrated by graphs of the Legendre polynomials shown in Fig. 8.1.

Property 8. This follows directly from Property 5 when the stated conditions are imposed, because increasing $p(x)$ will increase the numerator while leaving all other terms unchanged.

Property 9. No proof of this result is offered because it follows from the form of argument used to establish the upper bound property of the Rayleigh quotient given in (128).

Appendix 2 (Norm of $J_n(x)$)

The square of the norm of the Bessel function $J_n(j_{n,r}x/a)$ is the definite integral

$$\|J_n(j_{n,r}x/a)\|^2 = \int_0^a x[J_n(j_{n,r}x/a)]^2 dx = \frac{1}{2}a^2[J_{n+1}(j_{n,r})]^2,$$

and so the norm is

$$\|J_n(j_{n,r}x/a)\| = \frac{1}{\sqrt{2}}a[J_{n+1}(j_{n,r})]. \tag{147}$$

This result is most easily derived by considering the case $a = 1$, and then changing variables to obtain the foregoing more general result. Accordingly, we consider the two Bessel equations in Sturm–Liouville form,

$$[xu']' + (j_{n,r}^2 x - n^2/x)u = 0 \quad \text{and} \quad [xv']' + (k^2 x - n^2/x)v = 0,$$

defined on the interval $0 \leq x \leq 1$ with bounded solutions that satisfy the boundary conditions $u(1) = v(1) = 0$. These equations have the respective solutions $u(x) = J_n(j_{n,r}x)$ and $v(x) = J_n(kx)$.

Multiplying the first equation by u, the second by v, subtracting the second equation from the first, and integrating over the interval $0 \leq x \leq 1$ gives, after using Theorem 8.6 (ii) and the result $u'(x) = j_{n,r}J'(j_{n,r}x)$,

$$\int_0^1 xJ_n(j_{n,r}x)J_n(kx)dx = \frac{j_{n,r}J_n(k)J_n'(j_{n,r})}{k^2 - j_{n,r}^2}.$$

We now write this result as

$$\int_0^1 xJ_n(j_{n,r}x)J_n(kx)dx = \left(\frac{j_{n,r}}{k + j_{n,r}}\right)\left(\frac{J_n(k) - J_n(j_{n,r})}{k - j_{n,r}}\right)J_n'(j_{n,r}),$$

where the subtraction of $J_n(j_{n,r})$ in the bracketed term in the numerator leaves the result unchanged because $J_n(j_{n,r}) = 0$.

Taking the limit as $k \to j_{n,r}$, reduces this result to

$$\int_0^1 x[J_n(j_{n,r}x)]^2 dx = \frac{1}{2}[J_n'(j_{n,r})]^2, \quad r = 1, 2, \ldots.$$

It is inconvenient to work with $J_n'(j_{n,r})$, so we relate J_n to J_{n+1} by using recurrence relation (65)$'$:

$$xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x).$$

Setting $x = j_{n,r}$ causes this to simplify to $J_n'(j_{n,r}) = -J_{n+1}(j_{n,r})$, and so

$$\int_0^1 x[J_n(j_{n,r}x)]^2 dx = \frac{1}{2}[J_{n+1}(j_{n,r})]^2.$$

The more general result follows by making the change of variable $x = z/a$ and then replacing z by x.

EXERCISES 8.11

In Exercises 1 through 3 expand the given polynomials in terms of Legendre polynomials.

1. $4x^3 - 2x^2 + 1$.

2. $3x^3 + x^2 - 4x$.

3. $x^4 + 3x^2 + 2x$.

4. Represent x^2, x^3, and x^4 in terms of Legendre polynomials.

In Exercises 5 through 8 find the first four terms of the Fourier–Legendre expansions of the given functions. In each case graph the four term approximation to $f(x)$ and compare it with the graph of $f(x)$.

5. $f(x) = \begin{cases} 1, & -1 \le x \le 0 \\ x, & 0 < x \le 1. \end{cases}$

6. $f(x) = \begin{cases} 1+x, & -1 \le x \le 0 \\ 1-x, & 0 < x \le 1. \end{cases}$

7. $f(x) = \begin{cases} 0, & -1 \le x < -1/2 \\ 1, & -1/2 < x < 1/2 \\ 1/2, & 1/2 < x < 1. \end{cases}$

8. $f(x) = \begin{cases} -2x, & -1 \le x < 0 \\ x, & 0 \le x \le 1. \end{cases}$

9. Find the first four terms in the Fourier–Legendre expansion of e^x.

10. Find the first four terms in the Fourier–Legendre expansion of e^{-x}.

In Exercises 11 through 13 expand the given polynomials in terms of Chebyshev polynomials.

11. $3x^4 - 4x^2 - x$.

12. $4x^3 + x^2 - 3x + 1$.

13. $2x^4 - x^3 + x + 3$.

14. Represent x^2, x^3, and x^4 in terms of Chebyshev polynomials.

In Exercises 15 and 16 find the first four terms in the Fourier–Chebyshev expansion of the given function. In each case graph the four term approximation to $f(x)$ and compare it with the graph of $f(x)$.

15. $f(x) = \begin{cases} 2+x, & -1 < x < 0 \\ 3, & 0 < x < 1. \end{cases}$

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